

COLECCIÓN CIENCIA Y TECNOLOGÍA



# La metaética puesta a punto



**Guillermo Lariguet**

director

**Guillermo Lariguet**

**María Sol Yuan · Nicolás Alles**

compiladores

ediciones **UNL**



**UNIVERSIDAD  
NACIONAL DEL LITORAL**

 **ediciones UNL**

Consejo Asesor  
Colección Ciencia y Tecnología  
**Laura Cornaglia**  
**Miguel Irigoyen**  
**Luis Quevedo**  
**Alejandro Reyna**  
**Amorina Sánchez**  
**Ivana Tosti**  
**Alejandro Trombert**

Dirección editorial  
**Ivana Tosti**  
Coordinación editorial  
**María Alejandra Sadrán**  
Coordinación diseño  
**Alina Hill**  
Coordinación comercial  
**José Díaz**

Corrección  
**Félix Chávez**  
Diagramación interior y tapa  
**Laura Canterna**

© Ediciones UNL, 2023.

—  
Sugerencias y comentarios  
**editorial@unl.edu.ar**  
**www.unl.edu.ar/editorial**

La metaética puesta a punto / Nicolás Alles ...  
[et al.] ; compilación de Guillermo  
Lariguét ; María Sol Yuan ; Nicolás Alles ;  
dirigido por Guillermo Lariguét.- 1a ed.-  
Santa Fe : Ediciones UNL, 2023.  
Libro digital, PDF/A - (Ciencia y tecnología)

Archivo Digital: descarga y online  
ISBN 978-987-749-403-7

1. Filosofía General. 2. Ética. 3. Moral. I. Alles,  
Nicolás, comp. II. Lariguét, Guillermo, comp. III.  
Yuan, María Sol, comp.  
CDD 170.42

---

© Alles, Arena, Beade, Busdygan,  
Canclini, Chiloví, Daguerre, Flores,  
García Valverde, González Lagier,  
Kalpokas, Lariguét, Martínez Zorrilla,  
Mettini, Misseri, Oliveira,  
Paolicchi, Samamé,  
Truccone–Borgogno, Vercellone,  
Vidiella, Wagon, Yuan, 2023.

Revisión de originales: Joaquín Suárez



# La metaética puesta a punto

Guillermo Lariguet  
director

Guillermo Lariguet,  
María Sol Yuan y Nicolás Alles  
compiladores

Alles, Arena,  
Beade, Busdygan,  
Canclini, Chiloví,  
Daguerre, Flores,  
García Valverde,  
González Lagier,  
Kalpokas, Lariguet,  
Martínez Zorrilla,  
Mettini, Misseri, Oliveira,  
Paolicchi, Samamé,  
Truccone-Borgogno,  
Vercellone, Vidiella,  
Wagon, Yuan

**ediciones UNL**

CIENCIA Y TECNOLOGÍA

## Agradecimientos

Este libro ha sido posible gracias al interés de la decana de la Facultad de Humanidades y Ciencias de la Universidad Nacional del Litoral, Prof. Laura Tarabella, así como de la secretaria Académica de dicha institución, Mg. Bárbara Mántaras. Agradecemos también el apoyo de Ediciones UNL. Al minucioso empeño de Joaquín Suárez puesto en la revisión de los textos originales. Muchas gracias a los proyectos de investigación CAI+D UNL «Debates contemporáneos en torno a la justificación epistémica» (Dir. Eduardo Nicolás Giovannini) y «Teoría del discurso y Estado democrático de derecho» (Dir. Santiago Prono) por hacer posible la concreción de esta obra. Finalmente, queremos agradecer a cada autor/a quienes con sus contribuciones hicieron surgir este libro del que son parte.

# Índice

**Prólogo / 7**

## **I. Concepciones metaéticas: versiones clásicas y contemporáneas**

### **1. Naturalismo ético / 22**

*Martín Dağuerre*

### **2. Constructivismo moral / 50**

*Martín Oliveira*

### **3. Realismo moral / 63**

*Guillermo Lariguét*

### **4. Antirrealismo y Cuasirrealismo moral / 92**

*Federico Arena*

### **5. Escepticismo moral / 110**

*Leandro Paolicchi*

### **6. El relativismo metaético: argumentos a favor y en contra / 128**

*Lucas Misseri*

### **7. Particularismo y Universalismo moral / 141**

*Graciela Vidiella*

### **8. Intuiciones morales / 159**

*Guadalupe Mettini*

## **II. Problemas característicos**

### **9. Percepción moral: algunas clarificaciones y distinciones / 180**

*Daniel Kalpokas*

### **10. Argumentación moral / 199**

*Daniel Busdygan*

### **11. Experimentos mentales y el rol de la historia / 214**

*Santiago Truccone-Borgogno*

### **12. Desacuerdos morales / 229**

*David Martínez Zorrilla*

### **13. Metafísica y moral / 250**

*Samuele Chiloví*

### **14. Agencia moral, libertad de voluntad y determinismo / 272**

*Fernanda Flores*

### **15. El problema de la suerte moral / 285**

*Gustavo Beade*

**16. Neuroética, moral y el paso del «ser» al «deber ser» / 298**

*Daniel González Lagier*

**17. Motivación moral. El debate entre el internalismo y el externalismo / 312**

*Nicolás Alles*

**18. Akrasia o Debilidad de la voluntad / 335**

*Luciana Samamé*

### **III. Recientes proyecciones sobre problemas persistentes**

**19. Una lectura wittgensteiniana del campo de la ética / 354**

*María Sol Yuan*

**20. «Eso no es gracioso»: acerca de cómo los chistes misóginos ayudan a comprender la naturaleza de los valores / 372**

*Facundo García Valverde*

**21. Notas sobre el debate Kohlberg vs. Gilligan: de la diferencia de género a los desacuerdos metaéticos / 386**

*Adriana Vercellone*

**22. El ideal de pureza en la configuración de la identidad moral: conciencia e intersubjetividad desde el pensamiento de Hannah Arendt / 407**

*Rebeca Canclini*

**23. La metáfora del mal. En torno a la discusión entre Hannah Arendt y Gershom Scholem / 424**

*María Wagon*

**Sobre las autoras y los autores / 436**

## Prólogo

Guillermo Lariguet · María Sol Yuan · Nicolás Alles

La palabra «metaética», como casi cualquier palabra, especialmente filosófica, está «sobrecargada» de significados que, de no atenderse, implican el riesgo de que la filosofía conduzca cualquier debate conceptual como lo haría un auriga ciego. Entre estas sobrecargas está el hecho de que el término «metaética» suele asociarse a un momento arqueológico de la filosofía analítica cuando esta se identificaba con métodos propios del positivismo lógico. Por eso, cuando algunos filósofos y filósofas escuchan hablar de metaética, piensan de inmediato que se están evocando exclusivamente los avances, y prejuicios, de filósofos tan notables como Ayer, Stevenson o Carnap. Otra sobrecarga asociada con la anterior, pero que podemos deslindar, es aquella según la cual la metaética es un ejercicio «artificial», cuando no «artificioso», de análisis conceptual en el campo de la ética. Esta mirada ve en el ejercicio del metaético una distracción fútil, un pasatiempo con ínfulas de sofisticación.

Debido a las mencionadas sobrecargas semánticas, se ha dicho también, y esto añadiría una tercera significación detractora, que la metaética terminó de asesinar a la ética (léase: a la «ética normativa»). Que sin los enfoques característicos de estas éticas normativas clásicas, modernas o contemporáneas, la ética ya hubiera desaparecido de una currícula escolar colonizada —en forma agobiante— por las pretensiones imperiales de la metaética. Como eventual refuerzo de esta mirada kármica sobre la metaética, y del poder salvífico para el honor de los filósofos morales que vendrían a ejercer en exclusivo las éticas normativas, suele agregarse que las éticas «aplicadas», también, y de modo preponderante, tienen el potencial de redimir a aquellos filósofos de «gabinete» o «puros» —como los metaéticos—. A propósito de este último señalamiento, podemos recordar que Alasdair MacIntyre, en «¿Does applied ethics rest on a mistake?», nos interpelaba acerca de si el predicado «aplicada», adosado a la palabra ética, no descansaba, como reza el título de su artículo en inglés, en un error. Podemos interrogarnos, siguiendo a Stephen Toulmin, acerca de qué significa que la medicina, o la bioética, por ejemplo, le «salven» la vida a la ética o que, si no hay ética normativa, mejor «cerrar el negocio», porque la metaética es, como dijo alguna vez el querido Ernesto Garzón Valdés, «onanismo intelectual». Cuando nos preguntamos a dónde conducen estos calificativos conceptuales o des-calificativos no meramente teóricos sino más bien axiológicos, reinstalamos, *de jure*, la importancia de la metaética.

Contrariamente a Garzón Valdés, pensamos que la metaética y la ética normativa no tienen barreras lógicas que las dividan en forma infranqueable. Somos conscientes de que afirmar que no hay por fuerza tal brecha lógica insalvable es parte, probablemente, del problema filosófico consistente en saber cuál es el alcance de la metaética respecto de sus eventuales efectos prácticos. En el fondo, lo que acabamos de decir tiene cierto aire de familia con lo que pensaban filósofos tan diversos como Willard van Orman Quine, Karl Otto Apel, John Rawls, Ronald Dworkin o Ludwig Wittgenstein. Porque, ¿qué clase de barreras lógicas divisorias serían estas? ¿Qué diferencia tajante cabría estipular a la hora de analizar conceptos, por un lado y desear normar el mundo, por el otro? ¿Están la hermenéutica y la filosofía analítica realmente tan incomunicadas de modo que nuestros juicios normativos comprometidos o de primer orden están insalvablemente deslindados de sus presupuestos metaéticos más abstractos? Estas y otras preguntas, de tomarse con humilde seriedad, sugieren que la metaética está más viva que nunca y que, de no ser el caso, debiera resucitársela a como dé lugar.

Nuestra posición acerca del lugar y el rol de la metaética surge de la convicción de que ella no es quizás otra cosa que ética filosófica, como a veces se ha sugerido con valoración opuesta por cultores y denigradores; ella tiene un papel esclarecedor al que los filósofos *qua* filósofos no podemos, no debemos renunciar tan fácilmente. La razón es que los temas conceptuales, los problemas normativos, los métodos con que estudiamos filosóficamente las cosas, el soporte de nuestras afirmaciones o negaciones o la estructura y estatus de nuestras «teorías», generan una serie profusa de interrogantes conceptuales, metafísicos, epistémicos, semánticos, lógicos, normativos y metodológicos, que la metaética, por sus características connotadas de análisis de segundo orden, está preparada para atender de modo perspicuo.

La perspectiva que sostenemos respecto de la función y localización de la metaética no es incompatible, creemos, con discutir si de enunciados de segundo orden no podrían implicarse, *salva veritatis*, enunciados de primer orden, normativos o aplicados. No impide siquiera, por ímpetu lógico, escudriñar vasos comunicantes entre propuestas éticas normativas o aplicadas y sus correlativos ítems conceptuales, metafísicos, epistémicos, semánticos, lógicos, metodológicos. Si afirmásemos, por caso, que una virtud es una excelencia de carácter, o si dijésemos que los deberes morales son normativamente prioritarios y sistematizables, si sostuviésemos, también, que la ética debiera primar siempre sobre la política, o que la tortura debiera prohibirse de modo absoluto, o permitirse a veces, o que abortar durante los tres primeros meses está moralmente habilitado por una moral sana, o que la pobreza extrema es moralmente repugnante, o que el cambio climático requiere de regulaciones



estrictas, cualquiera sea la tesitura normativa que adoptásemos, nos tendríamos que preguntar, si somos filósofos, si somos honestos y sensibles, muchas cosas vinculadas. Por ejemplo, ¿cómo hay que entender una excelencia de carácter? ¿Forman un todo unitario y coherente las virtudes morales y las intelectuales? ¿Pensar en la prioridad de los deberes morales, supone desconocer otras variedades de deber o de operadores deónticos de estructura no reglada? ¿Qué suposiciones hacemos sobre el dolor cuando prohibimos la tortura o qué clase de factores especiales evaluaríamos de considerarla excepcional o trágicamente autorizada? ¿Al admitir el aborto durante el primer trimestre de embarazo suponemos algo así como la ausencia de una «mente fetal»? Si sostenemos que la ética debe primar sobre la política ¿significa que abrazamos la unidad del llamado razonamiento práctico? ¿Qué implicaría esta unidad? ¿No estamos dispuestos a admitir, como Maquiavelo sí hizo, auténticos conflictos entre, por ejemplo, deberes morales y políticos? Cuando hablamos con angustia de cambio climático y de la necesidad de regulaciones estrictas, ¿debemos suponer una concepción no antropocéntrica de nuestro mundo? Y si vehementemente luchamos contra la pobreza extrema, ¿qué factores, bienes, capacidades, necesidades, y bajo qué umbrales, consideraremos que la pobreza cuenta como extrema?

Los pocos ejemplos que hemos dado recién ilustran que la metaética, inclusive «deflacionada» respecto de sus antiguos prestigios, es inescapable. Y tal carácter inescapable no es el fruto del celo por la manutención de privilegios corporativos de la cofradía filosófica, sino por amor a la comprensión profunda y detallada de nuestros problemas éticos. En rigor, decir «detalles» no es abogar, necesariamente, por un análisis conceptual alambicado u oscuro hasta la extenuación o insensible a la relevancia práctica acuciante de los problemas prácticos de la ética.

## **Metaética puesta a punto**

¿Por qué hemos decidido titular esta obra con el giro no meramente retórico de «puesta a punto»? La respuesta será clara, esperamos, en lo que explicaremos a continuación.

Comencemos por señalar que a la fecha no hay constancia de publicación de un libro sobre Metaética en habla hispana que reúna similares características a las propuestas, a saber: que incluya temática y actualizaciones de la literatura de los últimos 20 años, sobre temas tanto clásicos (como libre albedrío, debilidad de la voluntad, realismo, antirrealismo, suerte moral) como más contemporáneos (experimentos mentales en ética, neuroética, género y razo-

namiento moral). En habla hispana pueden encontrarse con cierta dificultad ediciones antiguas —o traducciones de escasa calidad— de ediciones sobre el tópico en otros idiomas, o de ediciones en castellano difícilmente detectables en librerías de viejo o bibliotecas universitarias. Esta bibliografía pretérita, propia del positivismo lógico al que aludimos al comienzo de este prólogo, además, suele encontrarse enmarcada en una tradición más bien de corte marcadamente emotivista que ha sido sujeta a revisión por autores importantes como Simon Blackburn o Allan Gibbard, para poner solamente dos ejemplos destacados, aún no suficientemente traducidos al español. Sí puede encontrarse en abundancia, en español, y por lo comentado párrafos atrás, trabajos de ética normativa —y aplicada— pero bajo la forma de libros que tocan temas específicos, pero con despreocupación por la metaética, con algunas salvedades como los trabajos, por caso, para mencionar dos conspicuos ejemplos argentinos, de Osvaldo Guariglia y Graciela Vidiella, por una parte (*Breviario de ética*, Edhasa, Buenos Aires, 2011) o Ricardo Maliandi (*Ética: conceptos y problemas*. Biblos, Buenos Aires, 2004), por la otra.

En otros idiomas y traducidos al español podemos citar, solamente *ad exempla*, como referencias:

- 1) *Ordenando la ética* de Richard Hare, traducción de Joan Vergés Gifra, Ariel, Barcelona, 1997: ofrece una taxonomía de teorías éticas, fundamentalmente orientado a filósofos/as. Propone, además, un repaso definitivo de la posición ética fundamental del propio Hare, uno de los filósofos más influyentes en el área. La idea principal del libro es que plantear las cuestiones morales como cuestiones de hecho conduce inevitablemente al relativismo y nos ata a las culturas y a los lenguajes particulares. Las cuestiones morales son objetivas poniendo de relieve el carácter universalmente prescriptivo del lenguaje moral.
- 2) *El problema moral* de Michael Smith, traducción de R. Sánchez Brígido, Madrid, Marcial Pons, 2015: un rasgo dominante de la discusión en la filosofía moral contemporánea es la variedad de argumentos disponibles y el hecho de que, aparentemente, no hay a la vista un acuerdo sobre cuestiones morales fundamentales. Este libro presenta un diagnóstico de ese problema: aún no se ha lidiado adecuadamente con la tensión entre tres características centrales de nuestros juicios morales: la objetividad de esos juicios, su carácter práctico y un modo estándar de concebir la acción humana. El libro trasunta una discusión de humanismo versus kantismo, aunque no logra cubrir todos los problemas y perspectivas de nuestro libro.
- 3) *La Introducción a la filosofía moral* de James Rachels, traducida por G. Ortiz Millán, México, FCE, 2020: el libro reflexiona sobre problemas como el relativismo y el subjetivismo morales, la religión y su relación con la moralidad

- y el egoísmo ético y psicológico; asimismo, presenta algunas de las teorías normativas más importantes: la kantiana, el utilitarismo, las éticas de la virtud, la ética feminista y las teorías contractualistas. Sin embargo, la propuesta se centra más en la ética normativa y en menor medida en la metaética. Y, además, no toma en cuenta la discusión, por ejemplo, de si el marxismo no comporta también una ética normativa que deba tenerse en cuenta.
- 4) *El breviario de ética* de O. Guariglia y G. Vidiella, Buenos Aires, Edhasa, 2011: el libro propone una guía, especialmente dedicada a los estudiantes de ética en carreras filosóficas, para entender de qué hablamos cuando hablamos de ética, cuáles son las diferencias con la moral, cuáles fueron las mutaciones que tuvo a través de siglos y cuál es la manera de aplicarla a la vida cotidiana. Se trata más de ética normativa que de metaética, con excepción de la primera parte que es una introducción general, pero básica, al tema del conocimiento moral y cómo podría justificarse.
- 5) *Ética: conceptos y problemas* de R. Maliandi, Buenos Aires, Biblos, 2004: esta obra constituye una introducción general a la ética filosófica y su objetivo es contribuir a despertar algún interés por el estudio de los problemas éticos. El libro presenta, sin embargo, una visión más cercana a la ética del discurso y la ética material de los valores en esa peculiar versión «convergente» de Maliandi, pero, poco de metaética analítica, a la que le dedica solo un subapartado de menos de diez páginas.

Finalmente, en inglés, idioma característico de la metaética dominante, se puede hallar cuantiosa bibliografía, tanto compendios introductorios como específicos de algún enfoque particular. Entre estos últimos, pueden citarse, y de nuevo solamente a título de ejemplo, las siguientes referencias: La posición cuasirrealista del libro *Essays in Quasi-realism* de Simon Blackburn (Oxford U.P., 1993); el enfoque neo-emotivista de Allan Gibbard en *Wise choices, apt feelings: A theory of normative judgement* (Clarendon Press, Oxford, 1992); o la perspectiva realista moral de David Brink en *Moral Realism and the Foundations of Ethics* (Cambridge U.P., 1989). Con título «Metaética», y de carácter introductorio, pueden citarse: *An introduction to contemporary metaethics*, de A. Miller (Polity Press, 2003); *The Routledge Handbook of Metaethics* de Tristram Mc Person y David Plunkett (Routledge, 2019).

En resumen, como puede observarse a partir de los datos brindados, no hay a la fecha un libro en castellano, de esta amplitud temática, clásica y contemporánea, de esta diversidad de autores/as, y universidades.

## 2. Contenido y estructura de la obra

Como se viene sosteniendo, el libro presenta las discusiones actuales más relevantes sobre Metaética. Su objetivo general es analizar las distintas categorías epistémicas, metafísicas, psicológicas y semánticas que están en juego en los debates actuales en el campo de la Ética.

Un elemento destacado de la presente obra es que ofrece, hacia adentro de la Argentina, un enfoque «federal», y hacia afuera, una visión cosmopolita del mundo académico. Esto viene avalado por el hecho de que las contribuciones reúnen representantes destacados de varias universidades e institutos de investigación nacionales e internacionales. Entre los organismos nacionales, podemos nombrar siete universidades: Universidad Nacional del Litoral, Universidad Nacional de Córdoba, Universidad de Buenos Aires, Universidad Nacional de Quilmes, Universidad Nacional del Sur, Universidad Nacional de La Plata, Universidad Nacional de Mar del Plata. Los autores también radican en ocasiones su pertenencia en CONICET e institutos de doble dependencia: IHUCSO, CIF, KORN, SADAF, etc. Además, seis universidades internacionales se encuentran representadas a través de siete participaciones en el libro: Universidad Pompeu Fabra, Universidad de Barcelona (España), Universidad de Alicante (España), Universidad Oberta de Catalunya (España), Yachay Tech (Ecuador), Universidad de Graz (Austria), Universidad de California (Estados Unidos).

La amplia y variada participación comentada se complementa con la alta calidad de las contribuciones, las cuales recogen las categorías y debates actuales de la metaética, discutiendo críticamente, y a través de argumentos, sus empleos y potenciales ventajas. Entre los debates que podemos mencionar, se encuentran aquellos que giran en torno al realismo y el antirrealismo moral, el universalismo y el relativismo, la libertad, el libre albedrío y el determinismo moral, la discusión entre posiciones internalistas y externalistas en torno a la justificación de creencias/proposiciones morales; la naturaleza y contenido de las instancias que denominamos éticas y, finalmente, la reevaluación del campo y los límites propia de la disciplina ética, de la mano de consideraciones sobre género y la influencia de estereotipos en la validación de las instancias éticas.

Consideramos no estar errados si afirmamos que las contribuciones constituyen una herramienta importante para el estudio de la ética llevado a cabo por los estudiantes del grado y el posgrado, así como para investigadores académicos en el área más bien general de filosofía práctica, y especialmente de filosofía moral. Además, debido a que se trabaja con categorías y conceptos que se aplican a la ética, pero que también tienen incidencia en otras discipli-

nas teóricas y prácticas, las mismas son de utilidad en campos emparentados a las Ciencias Sociales, como son Ciencias Políticas, Sociología y Derecho.

En lo que concierne a la *estructura* que decidimos conferir a esta obra, hemos acordado dividirla en tres partes. La primera, de modo natural, introduce las principales versiones clásicas y contemporáneas de distintas concepciones metaéticas y agrupa las contribuciones de Martín Daguerre, Martín Oliveira, Guillermo Lariguet, Federico Arena, Leandro Paolicchi, Lucas Miseri, Graciela Vidiella y Guadalupe Mettini. La segunda parte del libro se introduce ya en problemas característicos de la disciplina, a saber: percepción moral, argumentación moral, experimentos mentales en ética, desacuerdos morales, las complejas raíces metafísicas de la moral, el problema clásico de la agencia moral y el libre albedrío versus determinismo. A su vez, en forma independiente pero conectada en parte con el anterior rango de temas, se halla una discusión del estatus de la suerte moral, el alcance de las neurociencias en ética bajo la condensación de la palabra «neuroética», el problema de la motivación moral acorde a enfoques de tipo internista y externista, respectivamente y, finalmente, un análisis distintivo de la *akrasia* respecto de la debilidad de la voluntad. Aportan sus trabajos esclarecedores sobre estos problemas Daniel Kalpokas, Daniel Busdygan, Santiago Truccone-Borgogno, David Martínez Zorrilla, Fernanda Flores, Gustavo Beade, Daniel González Lagier, Nicolás Alles y Luciana Samamé.

La tercera y última parte del libro lleva por título: «Recientes proyecciones sobre problemas persistentes». Desde luego, hay no solo problemas de ética filosófica persistentes como, por ejemplo, si la dimensión temporal (pasado-presente-futuro) debe afectar nuestros juicios normativos de primer orden (referidos a la justicia) sobre temas prácticos como el reconocimiento de derechos de colectivos originarios o cosas por el estilo. Además, existen temas relativamente novedosos como el cambio climático o el modo en que la injusticia epistémica afecta los procesos de percepción y deliberación morales. Sin embargo, no resulta posible abordar todas estas persistencias o recientes proyecciones filosóficas sobre temas sustantivos como los que acabamos de enunciar ejemplificativamente. Aquí hemos debido efectuar una selección más clásica que incorpora, antes que nada, una reflexión sobre el problema de delimitación de la ética. Aunque se trata de un indudable tema clásico, lo incluimos en este segmento por tratarse de un abordaje neo-wittgensteiniano que proyecta nueva luz sobre la cuestión. Esta misma política adoptamos al incluir en esta última parte el tema perenne de la naturaleza de los valores morales, ya que su autor aborda la cuestión desde una perspectiva original: la del alcance de chistes misóginos, vinculando, así, dos temas recurrentes, como son, los de la naturaleza del valor moral, por un lado, con la discusión

más bien normativa sobre los límites éticos de cierta clase de humor, por la otra. Asimismo, se compulsa el modo en que la llamada «perspectiva de género» permite complejizar los debates sobre el razonamiento moral, se discute el perenne tema de la identidad moral, pero con la novedad de una exploración metafórica del mismo y, al último, se introduce el también problema nodal del mal, pero revisitado bajo una óptica más amplia que la de la filosofía analítica, a través del debate entre Hannah Arendt y Gershom Sholem. Consecuentemente, por orden de aparición, intervienen en este examen teórico María Sol Yuan, Facundo García Valverde, Adriana Vercellone, Rebeca Canciani y María Wagon.

En lo que sigue, por si al lector le resultase útil, aunque esperamos no como dispensa de entusiasmarse por leer todo el libro en el orden que prefiera, presentamos una anticipación descriptiva y resumida del contenido del libro. Vale la pena insistir, antes del cierre, en que la metaética y su nueva puesta a punto es una empresa intelectual que los filósofos debemos seguir cultivando, como ocurre en la presente compilación. Incluso si hay que transformar el mundo como pedía Marx en su tesis onceava sobre Feuerbach, parece que, sin análisis conceptual, tal transformación pierde su esquema y debilita sus herramientas. A fin de cuentas, el propósito de este libro se encamina en el antiguo *dictum* socrático de que las vidas requieren exámenes o, ahora en los términos universalmente conocidos de Kant, en que las intuiciones sin conceptos son ciegas. Es verdad, sin embargo, que estos sin el contenido de la experiencia son siempre vacíos y por eso los filósofos que hacemos metaética debemos ser sensibles a los dolores normativos del mundo. Pero no menos cierto es el apotegma kantiano que levanta la pregunta de si siempre debemos acompañar cada representación —en este caso moral— por conceptos. Tal vez, después de todo, habrá o no contenidos «no conceptuales» como sostenía Gareth Evans, un tema que repercutirá sin dudas en tópicos como la percepción de particulares morales. Pero de nuevo: al evocar todo esto ¿no estamos dando vueltas a las tuercas de la metaética? Seguramente sí y por eso la metaética, pese a sus detractores, no está herrumbrosa; sigue viva y continúa siendo válida como herramienta auxiliar del pensamiento normativo y aplicativo en ética.

**I. Naturalismo ético** de Martín Daguerre (UNLP) ofrece un panorama de cuatro maneras de entender el naturalismo en el ámbito de la ética. Así, distinguirá: a) dos tipos de naturalismo metafísico: el naturalismo sintético no-reduccionista de Nicholas Sturgeon, representante del realismo de Cornell, y el naturalismo analítico reduccionista de Frank Jackson, representante del llamado Plan Canberra, y b) dos tipos de naturalismo me-

todo: el naturalismo sintético reduccionista de Peter Railton y el naturalismo evolutivo, representado por una posición no-cognitivistica, humano-darwiniana.

2. **Constructivismo moral** escrito por Martín Oliveira (UBA, CONICET) es un capítulo que reconstruye la posición que se conoce, en el ámbito de la filosofía práctica, como «constructivismo moral». A tal fin, se describe su surgimiento en la literatura especializada y su tesis básica. Además, se desarrolla una taxonomía de las distintas variantes del constructivismo, así como algunos de los ejemplos más importantes de ellas. Finalmente, se presenta una objeción al constructivismo y la mejor estrategia de respuesta a este problema.
3. **Realismo moral** es el capítulo perteneciente a Guillermo Lariguet (UNL, UNC, CONICET) en el cual se concentra en los elementos que deben adscribirse, al menos parcialmente, al realismo moral como categoría del pensamiento metaético. Esta posición puede resumirse de manera simple afirmando que hay aspectos del mundo en general, y del mundo moral en particular, que no dependen en un sentido importante de lo que pensemos, o deseemos. Este compromiso puede cobrar diferentes modalidades según la propuesta en la que se conjugue.
4. **Antirrealismo y cuasirrealismo moral** de Federico Arena (UNC, CONICET) aclara desde el inicio que no consistirá en una enumeración de las tesis que caracterizan al antirrealismo y al cuasirrealismo, sino que intentará mostrar qué tipo de preocupaciones llevaron a diferentes filósofos a sostener esas posiciones. Para ello, se comienza por la respuesta a la pregunta: ¿por qué el antirrealismo/cuasirrealismo? Luego siguen los principales argumentos en su defensa: la «superveniencia», la visión científica del mundo y, respecto del cuasirrealismo, la conservación del discurso moral. A continuación, se revisan algunas de las principales críticas: la dificultad para explicar el acto ilocutivo de expresar una evaluación moral, la imposibilidad de la distinción ética/metaética, entre otras.
5. **Escepticismo moral** redactado por Leandro Paolicchi (UNMP, CONICET) tiene como objetivo primordial presentar algunos de los argumentos más fuertes que circulan en el panorama contemporáneo acerca del escepticismo, así como de las formas más comunes de este pensamiento en el ámbito de la ética normativa y de la metaética actual. El resultado es un mapeo básico de los argumentos más consistentes de escepticismo moral en las discusiones de los últimos años. Las condiciones históricas y filosóficas que limitan las posibilidades de la ética normativa y su aplicación demarcan a su vez también las formas del escepticismo.

6. En **El relativismo metaético: argumentos a favor y en contra**, Lucas Misseri (Universidad de Alicante) ofrece un análisis del concepto de relativismo ético, partiendo de su oposición con el absolutismo ético y con conceptos afines como el objetivismo y el universalismo éticos. Asimismo, se ofrece una tipología de las distintas tesis que involucra, diferenciando entre el relativismo ético descriptivo, el ético normativo y el metaético.
7. **Particularismo y universalismo moral** desarrollado por Graciela Vidiella (UNL, UNLP) se propone subrayar que una de las cuestiones que no ha dejado de interesar a la ética occidental desde sus inicios es si el razonamiento y el juicio moral se orientan por principios o si se enfocan exclusivamente en los contextos y casos particulares. Los filósofos universalistas piensan que se determinan —o deberían hacerlo— por principios universales válidos para todas las situaciones. Bajo el concepto de particularismo, en cambio, se ha pretendido englobar posiciones bastante disímiles. El trabajo analiza críticamente esta disputa.
8. En **Intuiciones morales**, capítulo bocetado por Guadalupe Mettini (UNL), se propone un examen de la plausibilidad de la tesis de acuerdo con la cual las intuiciones desempeñan un rol fundamental en el razonamiento moral. De acuerdo con algunas conceptualizaciones las intuiciones morales compelen a los sujetos a formular un veredicto o determinan la inclinación a formular un juicio. Si la conclusión a la que se arriba como resultado de un razonamiento moral depende de la aceptación de ciertas premisas y las premisas deben su aceptabilidad a las intuiciones que las respaldan, el resultado de ciertos razonamientos morales depende de que los sujetos que juzgan tengan las intuiciones adecuadas. Por otro lado, las narrativas incluidas en algunas argumentaciones vinculadas a principios morales evocan escenarios que «gatillan» o disparan intuiciones que pueden socavar la universalidad de principios morales que se presumen universales. Por lo que, las intuiciones pueden desempeñar un rol relevante en el cuestionamiento de los principios que funcionan como premisas de los razonamientos morales.
9. **Percepción moral: algunas clarificaciones y distinciones** de Daniel Kalpokas (UNC, CONICET) reconoce, en general, que la experiencia perceptiva desempeña un papel esencial en la adquisición y justificación del conocimiento empírico. Se dice, por ejemplo, que las experiencias perceptivas justifican juicios y creencias empíricas. Ahora bien, suponiendo que haya algo así como «el conocimiento del mundo moral», cabe preguntar: ¿tiene la percepción un rol epistemológico semejante en la justificación de nuestros juicios y creencias morales?



- 10. Argumentación moral**, capítulo elaborado por Daniel Busdygan (UNQ) indaga en las características y complejidades de la argumentación moral. Él analiza los distintos niveles que pueden reconocerse en la argumentación moral y se propone presentar modos de esquematizar a los razonamientos morales. La complejidad relativa a la esquematización y a la evaluación de los razonamientos en general, se dirige a la clarificación de los tipos de premisas que componen al razonamiento en el marco monológico donde puede observar su necesidad, plausibilidad, revisabilidad o rebatibilidad. En el marco dialógico, se recuperan aspectos vinculados a la coherencia, eficacia y prudencia de la evaluación del argumento.
- 11. Experimentos mentales y el rol de la historia** de Santiago Truccone-Borgogno (Universidad de Graz) es un capítulo en el cual se sostiene que, de modo similar a las ciencias naturales donde se realizan experimentos en laboratorios con el fin de aislar elementos y analizar diferentes reacciones, los filósofos prácticos construyen diferentes clases de experimentos mentales con el objetivo de aislar propiedades relevantes que permiten testear principios y juicios normativos (morales y políticos). El trabajo reconstruye estos empleos y argumenta que el uso exclusivo de experimentos mentales es problemático en filosofía práctica.
- 12. Desacuerdos morales** redactado por David Martínez Zorrilla (Universidad Oberta de Catalunya) tiene por objeto intentar ofrecer una delimitación conceptual precisa de los diversos sentidos en los que puede hablarse de desacuerdos en materia moral, así como de sus principales implicaciones y su relación con la idea del error en el campo de la ética. Para ello es preciso analizar el concepto de desacuerdo en los ámbitos o dimensiones relacionados con el discurso moral, como la moral social, la ética normativa y la metaética, en sus principales versiones o corrientes de la discusión filosófica contemporánea. El análisis del concepto de desacuerdo y sus conexiones con el concepto de error permiten una visión más clara de las posibilidades de debate racional y respuesta correcta en el campo de la ética.
- 13. Metafísica y moral** de Samuele Chiloví (Universidad de California) sostiene que la metafísica es un campo de investigación amplio y variado, tanto por el alcance como por el rango de los temas que investiga. Por razones de espacio y relevancia, el autor solo se concentra en la parte de la metafísica que a veces se ha llamado «metafísica internivel» (*interlevel metaphysics*), y específicamente en una manera de desarrollarla que se apoya en la noción de fundamentación metafísica (*metaphysical grounding*). Como se verá, esta es la parte de la metafísica que tiene un impacto más directo para el estudio filosófico del derecho, lo que la vuelve especialmente relevante para los objetivos de este libro.

14. En **Agencia moral, libertad de voluntad y determinismo**, Fernanda Flores (UNL) mantiene la tesis según la cual el presupuesto fundamental de toda ética normativa es que los agentes morales gozamos de libre albedrío y somos responsables de nuestra conducta. Sin embargo, somos conscientes de que ciertos condicionamientos e influencias funcionan como atenuantes de la carga de responsabilidad que atribuimos a las personas y sus acciones. El capítulo explora esta tensión de la mano de varios filósofos referentes del tema.
15. **El problema de la suerte moral** es un capítulo en el cual Gustavo Beade (UBA) argumenta que el vínculo entre aquello que no controlamos y el modo en que juzgamos esas acciones constituye la llamada suerte moral, sencillamente ser alabado o reprobado en parte por la mera suerte. La suerte y el control, básicamente, son conceptos contrapuestos. De este modo, todo aquello que no está bajo mi control se encuentra librado a la suerte. El objetivo del capítulo es presentar algunos aspectos acerca de la suerte y también una introducción a la suerte moral.
16. **Neuroética, moral y el paso del «ser» al «deber ser»** de Daniel González Lagier (Universidad de Alicante) plantea que la neurociencia de la ética es el estudio de la conducta ética desde el punto de vista de las investigaciones sobre el funcionamiento del cerebro en los procesos cognitivos y de razonamiento. Se enfoca, por ejemplo, en la fundamentación de normas morales en emociones como la repugnancia, el análisis del papel de la oxitocina o de las «neuronas espejo» en nuestra conducta ética. El trabajo explora este campo metaético.
17. **Motivación moral. El debate entre el internalismo y el externalismo** es el capítulo de Nicolás Alles (UNL – UADER) en el cual se remarca que uno de los debates más importantes de la metaética contemporánea es aquel en torno a la motivación moral. ¿Qué nos motiva a actuar moralmente? ¿Cómo nos motivan los juicios morales? ¿Nos motivan de manera necesaria o contingente? En este trabajo Alles se centrará en analizar este debate desde la perspectiva del internalismo y del externalismo, dos posiciones que buscan responder estos interrogantes y otros al precisar la naturaleza de los juicios morales, su alcance y su relación con las creencias y otros estados conativos, como el deseo.
18. **Akrasia o debilidad de la voluntad** meditado por Luciana Samamé (Yachay Tech, Ecuador) parte de sostener que la cuestión de la *akrasia* —o debilidad de la voluntad— ha suscitado uno de los debates más arduos e inveterados en la historia de la filosofía moral. Es por ello por lo que el objetivo principal de este trabajo apunta a una reconstrucción conceptual del problema, lo más clara y despojada posible, sin desatender su dimensión histórica. Su formulación inicial aparece en el *Protágoras*, en donde Sócrates

tes establece que es imposible que alguien siga determinado curso de acción a sabiendas de que existe otro mejor. Con todo, filósofos subsiguientes han admitido la posibilidad de la *akrasia*, por lo que nos centraremos en los argumentos principales que se han dado a favor y en contra de dicho fenómeno, atendiendo también a otras discusiones conexas, tales como las referentes a la naturaleza del razonamiento práctico o al internalismo y externalismo de las razones.

**19. En Una lectura wittgensteiniana del campo de la ética.** María Sol Yuan (UNL, CONICET) analiza el dominio o campo propio de la ética a través de dos cruciales interrogantes a los que intenta responder: 1) ¿la ética versa sobre una dimensión individual de los seres humanos? y 2) ¿Hay lugar para los hechos morales? Para ello, se emplea cierta interpretación de dos posiciones claves de las *Investigaciones Filosóficas* de Ludwig Wittgenstein: su crítica al lenguaje ostensivo y el seguimiento de reglas que le sirven de corolario.

**20. «Eso no es gracioso»: acerca de cómo los chistes misóginos ayudan a comprender la naturaleza de los valores** es pensado por Facundo García Valverde (UBA, CONICET) como un asunto legítimamente metaético acerca de la naturaleza de los valores. La pregunta suele formularse así: ¿qué son los valores? Y continuarse de este modo: ¿son la proyección de nuestros deseos e intereses?, ¿son elementos de la fábrica del mundo que producen determinadas respuestas en el sujeto práctico?, ¿son análogos a propiedades primarias (forma, etc.) o a propiedades secundarias (color, etc.)?, ¿qué relación existe entre las emociones y los valores? Como se ve en este capítulo, el problema de la naturaleza de los valores es multifacético.

**21. Notas sobre el debate Kohlberg vs. Gilligan: de la diferencia de género a los desacuerdos metaéticos** escrito por Adriana Vercellone (UNC) es un capítulo en que se afirma que las personas aprendemos a pensar moralmente desde la infancia, gracias a la interacción con el entorno, la familia y la sociedad. Sin embargo, los estudios sobre razonamiento moral informan que las niñas y mujeres se encuentran en una etapa más baja del desarrollo cognitivo y moral. El capítulo explora algunas propuestas novedosas que reconocen la experiencia femenina en la definición de la identidad y del juicio moral.

**22. El ideal de pureza en la configuración de la identidad moral: conciencia e intersubjetividad desde el pensamiento de Hannah Arendt** de Rebeca Canclini (UNS) es un capítulo que tematiza la irrupción de construcciones identitarias regidas por ideales de pureza moral en el espacio público. Ella asume que este tipo de identidad se construye en la convergencia de ideales de limpieza y no-mezcla (santidad) que vinculan la conciencia individual con la ley moral sin referencia a un mundo ético compartido.

Se trata de construcciones identitarias que emergen en contextos de disolución social y, a la vez, aceleran la destrucción de los lazos comunitarios.

**23. La metáfora del mal. En torno a la discusión entre Hannah Arendt y Gershom Scholem** de María Wagon (UNS), sostiene, por último, pero de modo no menos relevante, que el tema del mal es dramáticamente neurálgico para la ética y se encuentra ampliamente extendido. El presente capítulo recoge una larga y compleja historia tanto en filosofía moral como política a través de la reconstrucción de un hito ilustrativo y relevante: la discusión que sobre el mismo entablaron Hannah Arendt y Gershom Scholem.

**I. Concepciones metaéticas:  
versiones clásicas  
y contemporáneas**

# 1. El naturalismo ético

*Martín Daguerre*

## Definición

Si bien el naturalismo ha sido comprendido de muchas maneras, podemos decir que se encuentra estrechamente relacionado con una posición metafísica y una posición metodológica.

Un naturalista metafísico entiende que las únicas propiedades que existen son las propiedades naturales, de manera que la realidad debe poder ser descrita en el lenguaje de las ciencias naturales. Llevado a la ética, decirse naturalista en este sentido es asumir la existencia de propiedades éticas y, dado que solo existen propiedades naturales, considerar que las propiedades éticas son propiedades naturales. Desde esta perspectiva, el lenguaje ético ofrece descripciones de la realidad. Decir que una persona es moralmente buena equivale a decir que posee alguna/s propiedad/es natural/es. Decir que la injusticia de las políticas aplicadas fue la causa del estallido social es, por tanto, ofrecer una explicación tan pertinente como cualquier otra explicación científica. Las propiedades éticas forman parte del mundo y, por lo tanto, intervienen en las cadenas causales de acontecimientos.

Para caracterizar al naturalismo metodológico, antes que identificarlo con la defensa de un método en particular para la investigación en filosofía, podemos asociarlo a la defensa de que no existen métodos *a priori* capaces de establecer verdades. Ninguna afirmación filosófica puede ser ajena a la crítica científica. De alguna manera, ciencia y filosofía deben ir de la mano.

Esta distinción es presentada con claridad por Peter Railton en *Naturalism and Prescriptivity*:

El naturalismo puede ser una doctrina acerca del método o de la sustancia. El *naturalismo metodológico* sostiene que la filosofía no cuenta con un método distintivo, *a priori*, capaz de producir verdades sustantivas que, en principio, no estén sujetas a ningún tipo de test empírico. En vez, un naturalista metodológico cree que la filosofía debe proceder *a posteriori*, en tándem con la investigación mayormente empírica llevada a cabo en las ciencias sociales y naturales —quizá como una parte particularmente abstracta y general de ella— la investigación mayormente empírica llevada a cabo en las ciencias naturales y sociales.

El *naturalismo sustantivo*, en cambio, no es en primera instancia una visión acerca de los métodos filosóficos, sino acerca de las conclusiones filosóficas. Un

naturalista sustantivo presenta una explicación filosófica de algún dominio del lenguaje o la práctica humanos que ofrece una interpretación de sus conceptos centrales en términos pasibles de investigación empírica. (Railton, 1989:155–156)

Un naturalista ético sustantivo, metafísico, no necesariamente será un naturalista metodológico. Como veremos, uno puede defender que las propiedades éticas son propiedades naturales, que los juicios éticos expresan creencias sobre estas propiedades, creencias que pueden ser verdaderas o falsas —de aquí que al naturalismo ético suele asociársele a una posición cognitivista—, y a la vez ofrecer un lugar al análisis filosófico que permite establecer verdades a priori.

Por su parte, el naturalista metodológico no tiene por qué defender un naturalismo ético sustantivo. Bien puede hacerlo, y esa será la posición del propio Railton, pero también un naturalista metodológico puede concluir que el no-cognitismo, por ejemplo, es la concepción metaética que resulta más convincente, dados los conocimientos científicos de nuestra época. En este sentido, podemos encontrarnos con naturalismos (metodológicos) éticos no-cognitivistas.

En lo que sigue, presentaré en primer lugar dos posiciones naturalistas sustantivas, metafísicas —una asumirá, también, la defensa de un naturalismo metodológico, y la otra recurrirá a análisis conceptuales a priori— y en segundo lugar dos posiciones naturalistas metodológicas: una que deriva en la defensa de un naturalismo sustantivo y otra que asume la defensa de una posición no-cognitivista.<sup>1</sup>

## **Naturalismo sustantivo**

Como vimos, el naturalismo ético sustantivo entiende que las propiedades éticas son propiedades naturales. La noción de propiedad natural debe tomarse en un sentido muy amplio, que incluye las propiedades abordadas tanto por las ciencias naturales, como por las ciencias sociales. Así, el tipo de receptores que posee una neurona, la intensidad del sufrimiento de una persona o el valor de una moneda son todas propiedades naturales.

---

1 El espacio acotado de este capítulo no permite desarrollar la posición de cada una de las cuatro líneas frente a distintos problemas metaéticos. Dentro del naturalismo encontraremos posiciones externalistas e internalistas, distintos modos de responder a la crítica de cometer la falacia naturalista, en sus diferentes versiones, diferentes explicaciones del desacuerdo moral, etc. El capítulo se limitará, en cambio, a presentar cuatro modos naturalistas de entender la tarea de la ética como disciplina.

Cabe agregar, como rasgo general del naturalismo ético sustantivo, la consideración de que las propiedades morales supervienen a las propiedades naturales no-morales, esto es, si dos cosas tienen las mismas propiedades naturales, tendrán las mismas propiedades morales. En el caso de que dos cosas tengan diferentes propiedades morales, deberán tener también diferentes propiedades naturales. No tiene sentido evaluar de manera diferente dos acciones con las mismas propiedades naturales, en las mismas circunstancias naturales.

Veamos, entonces, dos desarrollos diferentes de esta posición: un naturalismo sintético no reduccionista, tal como es presentado por Nicholas Sturgeon,<sup>2</sup> y un naturalismo analítico reduccionista, tal como es presentado por Frank Jackson.<sup>3</sup>

### El naturalismo sintético no reduccionista: Nicholas Sturgeon

Según Sturgeon, el naturalismo ético sostiene

(a) que propiedades éticas tales como la bondad de las personas, de los rasgos de carácter y de otras cosas, y como la corrección o incorrección de acciones, son propiedades naturales del mismo tipo general que las propiedades investigadas por las ciencias, y (b) que tienen que ser investigadas del mismo modo general en que investigamos a estas últimas. (Sturgeon, 2006a:92)<sup>4</sup>

Para poner en contexto al naturalismo ético sustantivo vale destacar que el inicio del siglo xx se vio marcado por una crítica de G.E. Moore (1903) al naturalismo que, para muchos, fue decisiva. Si bien la crítica fue muy influyente, el no-naturalismo defendido por Moore no tuvo la misma suerte, fundamentalmente por la desconfianza que generaba su postulación de propiedades *sui generis*. Por lo tanto, la posición que fue ganando terreno fue el no-cognitivism, en la medida en que no postulaba propiedades extrañas, a la vez que no se veía afectado por la crítica de Moore. El naturalismo se encontró, entonces, a la defensiva. Necesitaba ofrecer una respuesta a la crítica de Moore, que volviese a esta corriente una rival seria del no-cognitivism.

---

2 Nicholas Sturgeon es un representante de lo que se conoce como *el realismo de Cornell*.

3 Frank Jackson es uno de los principales impulsores de lo que se conoce como *el Plan Canberra*.

4 Hay una prioridad de a) sobre b): «He definido el naturalismo ético primariamente como una doctrina metafísica (acerca de qué tipo de hechos son los hechos morales), aunque con un nada sorprendente corolario epistemológico (que, dado que los hechos morales son hechos naturales, podemos adquirir conocimiento sobre ellos prácticamente del mismo modo en que adquirimos conocimiento sobre otros hechos naturales)». (Sturgeon, 2006a:93). Como veremos, Railton invertirá esta relación entre el naturalismo sustantivo y el naturalismo metodológico.



Sturgeon presenta del siguiente modo la objeción de Moore: dos propiedades son idénticas si los términos con los que se enuncia la identidad son sinónimos. Ahora bien, dado que ningún término que refiera a una propiedad natural (por ejemplo, *placentero*) tiene el mismo significado que un término ético (por ejemplo, *bueno*), ningún término ético puede referir a una propiedad natural. El naturalismo ético sintético<sup>5</sup> (tanto el reduccionista como el no-reduccionista) estará de acuerdo con Moore en que ningún análisis del significado de *bueno* determinará su sinonimia con un término que refiera a una propiedad natural; sin embargo, ello no quiere decir que *bueno* no refiera a una propiedad natural. En principio, es claro que el agua y la unión molecular  $H_2O$  son propiedades idénticas; aun así, el significado de ambos términos es diferente. También es claro que ningún análisis del significado de *agua* en el siglo XVIII hubiese arrojado que significaba lo mismo que esa unión molecular entre hidrógeno y oxígeno. Ahora bien, hoy por hoy, se puede definir correctamente a *agua* como  $H_2O$ . La expresión de esta identidad es verdadera, pero producto de una investigación científica y siempre sujeta a revisión.<sup>6</sup> Sturgeon aceptará, entonces, que es factible lograr establecer una identidad entre la propiedad de ser bueno y otras propiedades naturales. De hecho, da por supuesto que toda propiedad ética superviene sobre propiedades físicas.

Pero esto no implica que estemos en condiciones de decir con términos físicos todo lo que podemos decir con términos éticos. Los términos éticos refieren a propiedades naturales, que supervienen sobre propiedades físicas, pero a las que posiblemente no podamos referirnos con términos físicos (no contamos con la cantidad suficiente de términos físicos como para hacer referencia a cualquier propiedad natural). Los términos biológicos también re-

---

5 El naturalista sintético considerará que la determinación de qué propiedad natural es una propiedad ética no puede hacerse mediante un análisis del significado de los términos éticos, sino que requiere de una investigación científica.

6 El planteo sigue la línea abierta por los trabajos de H. Putnam y S. Kripke en torno a la referencia. R. Boyd presenta la teoría de la referencia del naturalismo sintético de la siguiente manera: «un término *t* refiere a una clase (propiedad, relación, etc.) *k* sólo en caso de que existan mecanismos causales cuya tendencia es provocar, con el tiempo, que lo que es predicado del término *t* sea aproximadamente verdadero de *k* (...) Tales mecanismos típicamente incluirán la existencia de procedimientos aproximadamente precisos para reconocer miembros o ejemplos de *k* (al menos para los casos fáciles) y que gobiernan de una manera relevante el uso de *t*, la transmisión social de ciertas creencias aproximadamente (de una manera relevante) verdaderas con respecto a *k*, formuladas como afirmaciones sobre *t* (...), un patrón de deferencia hacia los expertos sobre *k* en relación con el uso de *t*, etc. (...) Cuando prevalecen relaciones de este tipo, podemos pensar que las propiedades de *k* están regulando el uso de *t* (vía tales relaciones causales), y podemos pensar que lo que se dice usando *t* nos ofrece un acceso epistémico, socialmente coordinado, a *k*; *t* refiere a *k* (...) sólo en caso de que el uso socialmente coordinado de *t* ofrezca un acceso epistémico significativo a *k*, y no a otras clases (propiedades, etc.)» (Boyd, 1988:195).

fieren a propiedades naturales que, en última instancia, supervienen sobre propiedades físicas, pero el no contar con términos físicos para decir lo que podemos decir con términos biológicos no inhabilita a los términos biológicos para formar parte de explicaciones científicas. La búsqueda de definiciones reductivas es, para Sturgeon, inviable e innecesaria.

De esta manera se eluden los problemas semánticos, y el problema a resolver pasa a ser el epistemológico. Si consideramos que las propiedades éticas son propiedades naturales, pero no podemos ofrecer una definición reductiva, física, de las mismas, ¿cómo podemos conocer algo sobre ellas? ¿Por qué pensar, incluso, que son propiedades naturales? La estrategia de Sturgeon es colocar a las propiedades éticas en una red causal. Si esto es posible (y damos por supuesto que no existen propiedades sobrenaturales que intervienen en la red causal, en un plano natural), tendría sentido sostener que las propiedades éticas son propiedades naturales, y ubicar una propiedad en la red nos aportaría conocimiento sobre la propiedad misma, en la medida en que muestra cuáles son sus causas y/o consecuencias.<sup>7</sup>

Tal postura se enfrenta al desafío presentado por Gilbert Harman. Harman inicia *The Nature of Morality* con la siguiente pregunta: «¿Los principios morales pueden ser testeados y confirmados de la misma manera que los principios científicos?» (1977:3). Más precisamente, la pregunta es: ¿puede la observación jugar en la ética el papel que juega en la ciencia?

Harman no parte de una concepción ingenua de la observación. Él entiende que no hay observaciones puras. La observación implica formar una creencia como resultado directo de ver algo, y la creencia que nos formemos dependerá de la teoría que hayamos aceptado. Esto ocurre tanto en la ciencia, como en la ética: dada la teoría aceptada, frente a ciertos hechos nos formaremos inmediatamente una creencia, sin necesidad de llevar a cabo un proceso deliberativo. Tomando esta concepción de «observación», incluso podemos aceptar que las teorías éticas pueden confirmarse o refutarse apelando a la observación. Así, si diseñamos una situación que nos dispare la creencia inmediata de que lo que allí ocurre está mal, pero según nuestra teoría está bien,

---

7 R. Boyd ofrece una analogía entre la investigación sobre la bondad moral y la investigación sobre la salud. Es claro que la salud es una propiedad natural, pero no podemos ofrecer una definición reductiva de la misma. Sí podemos, en cambio, ir descubriendo conexiones causales entre la salud y diversas propiedades naturales (desde condiciones ambientales y sociales, a hábitos personales), que nos permitirán ir aumentando nuestro conocimiento sobre la salud. Una teoría ética debería contar con cierta noción de bondad moral y comenzar a analizar cómo la misma forma parte de cadenas causales junto con otros eventos morales y no-morales (como la disposición a actuar justamente, un sentimiento de autorrealización o el logro de ciertos placeres). Ver Boyd (1988).

deberemos o bien corregir la teoría, o bien aceptar que nuestra observación estaba sesgada.<sup>8</sup>

Sin embargo, dirá Harman, esto no quiere decir que estemos contrastando la teoría con la realidad. Hay una diferencia importante entre la observación en física y la observación en ética, aun entendiendo que la observación siempre depende de la teoría. En física la observación juega el papel de evidencia de una teoría, mientras que en ética no parece poder jugar ese rol.

Una teoría científica postulará hechos a los que podemos acceder por medio de la percepción de sus consecuencias. Tomemos el caso, presentado por el propio Harman, de un físico que al ver una estela de vapor en una cámara de niebla, inmediatamente piensa: «Allí va un protón». ¿Por qué hace esta observación? Porque su teoría postula la existencia de ese hecho como la mejor explicación de la presencia del vapor.

Pasemos, ahora, al ejemplo ético que presenta Harman: una persona, al doblar en una esquina, ve a un grupo de chicos que echa combustible sobre un gato y lo prende fuego, e inmediatamente piensa: «Eso es incorrecto». ¿Por qué hace esta observación? No es porque su teoría, al postular un hecho moral, contribuya a explicar lo que los chicos hicieron; esto es, no cree que sea incorrecto porque postular la existencia de esta propiedad es lo que mejor explica la acción de los chicos. La observación bien pudo haberla hecho por la teoría ética que acepta, que incluye la idea de que es incorrecto dañar a los animales, pero no es el hecho de que sea incorrecto dañar a los animales lo que llevó a los chicos a prender fuego al gato. Si la observación es verdadera o falsa, esto es, si es verdadero o falso que es incorrecto lo que los chicos están haciendo, parece irrelevante para explicar sus actos (y no era irrelevante que estuviese o no presente el protón para explicar la estela de vapor).

Como las teorías éticas no contribuyen a explicar hecho alguno, ninguna observación será evidencia a favor o en contra de las mismas.

Harman concluye:

---

8 Sin explicitarlo, Harman estaría haciendo referencia al proceso de mutuo ajuste entre principios y juicios, que busca lograr un equilibrio reflexivo. La aplicación de este procedimiento a la filosofía moral había sido defendida por John Rawls (1971). Harman considera que esto es un error, mientras que el realismo de Cornell reivindicará tal procedimiento, pero desde una posición realista, ausente en el planteo de Rawls. Los realistas de Cornell consideran que, si partimos de una teoría ética aproximadamente verdadera, el procedimiento de mutuo ajuste nos permitirá realizar nuevos descubrimientos. Entienden, a su vez, que este mismo procedimiento es el que se aplica en toda ciencia natural. Para la crítica de Harman, ver Harman (1986). Para una interpretación realista del equilibrio reflexivo, ver Boyd (1988).

En el caso científico, el hecho de que hagas esa observación es en sí mismo evidencia a favor de la teoría física porque la teoría física explica el protón, que explica la estela, que explica tu observación. En el caso moral, el hecho de que hagas tu observación no parece ser evidencia a favor del principio moral relevante porque ese principio no parece ayudar a explicar tu observación. La cadena explicativa desde el principio a la observación parece estar rota en el caso de la moralidad. (1977:8)

Sturgeon sostiene que no es cierto que esa cadena explicativa no se dé en el terreno de la moralidad. Aplica su análisis a un ejemplo centrado en la figura de Hitler. Supongamos que alguien, al ver un conjunto de acciones llevadas a cabo por Hitler, cree automáticamente que Hitler es una persona moralmente depravada. ¿Esta observación no podría explicarse por el hecho de que Hitler realmente es una persona moralmente depravada? Según el planteo de Harman, la creencia en el carácter moralmente depravado de Hitler sería irrelevante de cara a la explicación de sus acciones, si fuese indiferente para ello que la creencia sea verdadera o falsa. Pero apliquemos el siguiente test contrafáctico: si el hecho moral observado fuese falso, esto es, si Hitler no hubiese tenido un carácter moralmente depravado, ¿hubiese hecho lo que hizo? Difícilmente, razón por la cual cabe pensar que el carácter moral de Hitler efectivamente explica su comportamiento y, por lo tanto, explica la observación, esto es, la creencia de que tiene un carácter moralmente depravado.<sup>9</sup>

Podría objetarse que decimos que el acto es incorrecto, o que Hitler tiene un carácter moralmente depravado, desde una teoría ética en particular. Si en realidad no fuese incorrecto quemar al gato, igual pensaríamos lo mismo, porque nuestra sensibilidad y la teoría que aceptamos nos llevan a observar que el acto es incorrecto. Sturgeon señala que nuestra teoría moral debería estar radicalmente equivocada para que sea falso que un acto como el de incendiar un gato fuese incorrecto, o para que sea falso que un carácter como el de Hitler sea moralmente depravado. En ese caso, efectivamente se daría lo

---

9 En el ejemplo sobre Hitler, el hecho moral de que Hitler fue moralmente depravado explica la creencia de que fue moralmente depravado, porque el carácter moral de Hitler es lo que mejor explica sus acciones. La aplicación del enfoque de Sturgeon al ejemplo de Harman es menos transparente. Para aplicar el test contrafáctico, deberíamos preguntarnos qué hubiésemos observado si el hecho moral fuese falso. Si fuese falso que el acto de los chicos es incorrecto, ¿aun así creeríamos que es incorrecto? Dada la superveniencia de las propiedades morales sobre las naturales, para que fuese falso que el acto es incorrecto, deberían modificarse rasgos de la situación. Pero si modificamos rasgos de la situación, ya no podemos afirmar con seguridad que aun así consideraríamos que el acto es incorrecto. De manera que ciertos rasgos naturales de la situación son los que podrían explicar la observación. Para el desarrollo de la respuesta a Harman, ver Sturgeon (1986).

que Harman pretende sostener: si nuestra teoría moral fuese totalmente incorrecta y los hechos no-morales fuesen los destacados, entonces sacaríamos las mismas conclusiones que si la teoría moral fuese correcta. Ahora bien, dirá, nada diferente ocurre en la física. Si no hay un protón, pero se presenta la estela de vapor, el físico igualmente creerá que hay un protón. Y ello porque cree en una teoría muy errónea. Si la teoría es errónea y los hechos son los mismos, entonces el físico, dado que acepta la teoría, sacará las mismas conclusiones que saca.

Si bien Sturgeon pretende haber superado el desafío de Harman, quizá sea más apropiado entender su respuesta como un contra-ataque. Harman veía a la ética (en caso de no poder ofrecer definiciones reductivas) como limitada a un equilibrio reflexivo, mientras que la ciencia coteja sus teorías contra el mundo. Sturgeon sostendrá que, en realidad, todas las ciencias se limitan a usar el equilibrio reflexivo, razón por la cual la ética no tiene un problema particular.

En ética haríamos lo siguiente:

comenzamos con una amplia variedad de visiones evaluativas sobre muchos temas y las refinamos haciéndolas jugar unas contra otras y contra la evidencia empírica, en una búsqueda de lo que John Rawls llama equilibrio reflexivo (...) Las observaciones evaluativas y, más generalmente, los casos en los cuales (como sostengo) los hechos o propiedades evaluativas juegan un rol en la causación de nuestras creencias evaluativas, son un elemento en esta mezcla, pero no tienen un rol privilegiado. Precisamente porque están sesgadas por la teoría, pueden ser incorrectas cuando se basan en teorías evaluativas erróneas, y podemos siempre, en principio, cuestionar las teorías. Pero cuando el procedimiento completo está funcionando apropiadamente, ellas, y las creencias de fondo sobre las que descansan, pueden contar aun así como conocimiento. (Sturgeon, 2006b:254)

Nada diferente se hace en ciencia: «El pensamiento científico consiste en ajustar unas contra otras, y contra la evidencia empírica, una amplia variedad de visiones teóricas que tentativamente ya sostenemos, en búsqueda de un equilibrio reflexivo» (255).

En conclusión, partiendo de teorías evaluativas que no son radicalmente falsas, que incorporan términos que refieren a propiedades naturales,<sup>10</sup> es factible ir avanzando, por medio del equilibrio reflexivo, en nuestro conocimiento de las propiedades éticas, principalmente por medio de la determinación de sus causas y consecuencias.

---

10 Téngase en cuenta que el vínculo entre un término y su referencia es causal, tal como se explicitó en la nota 6.

## El naturalismo analítico: Frank Jackson

Al igual que Sturgeon, Jackson asocia al naturalismo ético con el naturalismo metafísico:

A grandes rasgos, el naturalismo ético es la doctrina que sostiene que las propiedades éticas son propiedades naturales. ¿Qué es una propiedad natural? Me refiero a las clases de propiedades discutidas en las ciencias naturales, entendiendo a las ciencias naturales de manera inclusiva (...) En la práctica esto significa que las propiedades naturales son las clases de propiedades que los teóricos tienen en mente cuando insisten en que lo moral superviene sobre lo no-moral, que lo normativo superviene sobre lo no-normativo. Las propiedades naturales son las bases de superveniencia para aquellas dos famosas afirmaciones de superveniencia. El naturalismo es, por lo tanto, una doctrina metafísica, tanto como ética. (Jackson, 2012:70)

A diferencia de Sturgeon, sin embargo, entenderá que la aceptación de la superveniencia de la ética sobre las propiedades naturales nos lleva a buscar una definición de los términos morales utilizando términos descriptivos no-morales. Su idea es que,

si las propiedades éticas son patrones dentro de lo no-moral, todo lo que se puede decir acerca de la naturaleza ética puede decirse usando las palabras correctas para aquellos patrones no-morales. Por lo tanto, existen oraciones formuladas enteramente en términos no-morales que hacen exactamente el trabajo que las oraciones éticas tienen en la imagen cognitivista. (Jackson, 2005:101)

Es fundamental para el naturalismo ético sustantivo dar cuenta de esta relación, esto es, explicar de qué manera los términos éticos pueden reducirse a un lenguaje descriptivo, y Jackson sostiene que esa relación debe establecerse a priori.<sup>11</sup>

Jackson compara el significado de los términos éticos con el significado de «densidad». La densidad de los cuerpos depende de su masa y volumen, pero no equivale ni a masa, ni a volumen. Hablar de densidad equivale a reunir los datos bajo cierto concepto. Al hablar de la densidad de un objeto, estamos diciendo algo sobre cómo son las cosas, pero esto no quita que la

---

11 De esta manera, Jackson considera que su posición consiste en un *descriptivismo metafísico analítico*. Posiciones como la de Sturgeon entrarían en lo que Jackson llama *descriptivismo metafísico ontológico*.

densidad pueda reducirse a propiedades naturales sobre las que superviene. Para Jackson,

deberíamos decir lo mismo en relación con el vocabulario ético: es parte implícita (si fuera explícita, la cuestión no sería filosóficamente controvertida) de nuestra comprensión de los términos y oraciones éticos que ellos sirven para marcar distinciones entre las maneras descriptivas de cómo son las cosas. (1998:125)

En la medida en que usamos los conceptos éticos para ofrecer información sobre cómo son las cosas, tiene que haber cierto consenso sobre su significado; de otro modo, no entenderíamos qué nos están comunicando mediante los conceptos éticos. Él compara nuestro dominio del lenguaje ético con nuestro dominio de un mapa. ¿Por qué nos resulta de utilidad que nos presenten un mapa cuando estamos perdidos? Porque el mapa representa el mundo de una manera que nosotros conocemos; esto es, el mapa nos dice algo sobre cómo es el mundo y, además, nosotros sabemos cómo leer un mapa. Del mismo modo, en ciertas circunstancias es pertinente emitir un juicio ético, o nos puede ser de utilidad escuchar uno, porque los mismos transmiten algo sobre cómo es el mundo, y nosotros sabemos interpretar lo que esos juicios nos dicen del mundo.

Jackson considera que el significado de cada término moral viene dado por su lugar en una red de conceptos interconectados. De esta manera, la comprensión de un término puede ir de la mano con el desconocimiento de la referencia, pero en cierto sentido. Así, comprendemos el significado de «la mujer más alta», pero difícilmente sepamos quién es la mujer más alta. Sin embargo, en otro sentido, comprender el término implica conocer la referencia: sabemos qué propiedad tiene que tener alguien para ser «la mujer más alta». Por lo tanto, cuando alguien dice que «María es la mujer más alta», sabemos qué es lo que nos está diciendo en relación con cómo son las cosas.<sup>12</sup> En este sentido, conocer el significado de un término equivale a conocer sus condiciones de verdad.

---

12 De esta manera se separa de la teoría de la referencia adoptada por los naturalistas sintéticos.

Según Jackson, tal teoría nos dice qué es lo que alguien (al usar términos éticos) está diciendo sobre cómo son las cosas, pero lo que dice depende de cuestiones que escapan a la mera comprensión de las palabras. Esto nos dejaría ante dos opciones, ambas poco atractivas en relación con los términos éticos: a) o bien nos encontramos como antes de conocerse la estructura química del agua y, por lo tanto, cuando utilizamos términos éticos hacemos referencia a cómo es el mundo, pero no sabemos, en realidad, cómo es. En tal caso, lo que correspondería hacer es dejar de usar los términos éticos; b) o bien nos encontramos como después de conocerse la estructura química del agua, pero ello implicaría que ya estamos ante una definición naturalista consensuada de, por ejemplo, «correcto», lo cual parece muy lejano a la realidad.

Esto abre el juego a dos tareas: a) una tarea analítica, que determine el significado de los términos éticos analizando la red de conceptos interconectados; y b) una tarea empírica que determine qué propiedad natural en particular satisface las condiciones de verdad del término moral.

El punto de partida del análisis será la moralidad *folk*, que nos da la base de consenso que nos permite debatir en ética. La moralidad *folk* es:

la red de opiniones, intuiciones, principios y conceptos morales cuyo dominio es parte integral de tener un sentido de lo que es correcto e incorrecto, y de ser capaz de participar en un debate significativo sobre lo que debe hacerse. Podemos pensarla como a la psicología *folk*, teniendo una naturaleza tripartita: como la psicología *folk*, contiene cláusulas de entrada, cláusulas con rol interno y cláusulas de salida. Las cláusulas de entrada de la moralidad *folk* nos dicen qué clases de situaciones descritas en términos no-morales, descriptivos, justifican qué clases de descripciones en términos éticos: si un acto es un asesinato intencional, entonces normalmente es incorrecto; el dolor es malo; «yo corto, vos elegís» es un procedimiento justo; etc. Las cláusulas con rol interno de la moralidad *folk* articulan las interconexiones entre cuestiones descritas en lenguaje ético, normativo: es más probable que las personas valientes hagan lo que es correcto, que las personas cobardes; la mejor opción es la opción correcta; los derechos imponen deberes de respeto; etc. Las cláusulas de salida de la moralidad *folk* nos llevan desde los juicios éticos a juicios sobre las motivaciones y, por consiguiente, sobre la conducta: el juicio de que un acto es correcto normalmente va acompañado por al menos algún deseo de realizar el acto en cuestión; la comprensión de que un acto sería deshonesto normalmente disuade a un agente de realizarlo; las propiedades que hacen a algo bueno son las propiedades hacia las que típicamente tenemos alguna especie de disposición favorable, etc. (Jackson, 1998:130-131)

El análisis de los significados de los términos éticos será, entonces, un trabajo a priori.<sup>13</sup> No tiene sentido, para Jackson, que diseñemos experimentos

---

13 La estrategia de trabajo es la que ha venido a denominarse *el Plan Canberra*. Si queremos determinar el significado de *correcto*, deberemos establecer todas sus conexiones de sentido común (*platitudes*) en la red. Tomamos la conjunción de todas las conexiones y a cada *conyunto* lo expresamos como una propiedad. Luego, en el lugar de la propiedad ponemos una variable y la cuantificamos existencialmente. Ahora tenemos que *correcto* es igual a una fórmula totalmente descriptiva, sin términos normativos, lo cual da cumplimiento a la tesis de superveniencia. Finalmente, deberemos determinar empíricamente qué propiedad natural, en lugar de la variable correspondiente, vuelve verdadera a la fórmula.



para determinar si lo que es bueno tiene relación con lo que es correcto o si existe una relación entre derechos y deberes.<sup>14</sup>

Ahora, si bien es cierto que comprender un concepto ético equivale a comprender su lugar en la red, también es cierto que la interpretación precisa de cada término dentro de la red es discutible. Así, filósofos consecuencialistas y deontológicos discutirán sobre el modo de interpretar la relación entre bueno y correcto, pero discuten sobre ello porque están de acuerdo en que alguna relación existe entre ambos términos. Incluso uno podría sostener que no existe ninguna relación entre los mismos, pero debería ofrecer una extensa argumentación para defenderlo, y ¿por qué debería hacerlo, si no es porque hay una red consensuada contra la que testear las posiciones éticas? Del mismo modo, uno puede ser externalista o internalista con respecto a la relación entre juicios morales y acciones, pero de fondo hay un acuerdo en que ambos ítems están relacionados.

Durante el desarrollo mismo del análisis puede ocurrir que la misma red se modifique:

la visión de la red permite que haya una comunión sustantiva en medio de esta diversidad. Desde esta visión, la meta última de la teoría ética es dar el mayor sentido a la red, lo cual puede implicar una modificación de la red. La red no es como un conjunto fijo de ecuaciones matemáticas simultáneas esperando ser resuelto. Podemos decir que la red es un trabajo en desarrollo, evolucionando por efecto de la reflexión crítica, pero posiblemente (*posiblemente*) convergiendo en lo que en otro lugar he llamado moralidad *folk* madura.<sup>15</sup> (Jackson, 2012:87)

Dicho de otra manera, el trabajo es el siguiente. Tomamos un concepto ético, por ejemplo «bueno», y establecemos todas las conexiones que tiene con los otros términos. De esta manera estableceremos su significado. Una

---

14 Si alguien no ve la conexión, es porque directamente no entiende de qué estamos hablando. En *On ethical naturalism and the philosophy of language* lo presenta del siguiente modo: «Pensemos en lo que uno diría a alguien que no entendió cosas como: que hay una estrecha conexión entre la bondad y lo correcto, entre derechos y deberes, entre ser virtuoso y hacer lo correcto, entre conceder un derecho a hacer x y permitir que se haga x, y entre el carácter apropiado del castigo y que el castigado sea culpable, o que la equidad importa para lo que debe ser hecho, que la moralidad se conecta íntimamente con la acción, que hay una conexión no-accidental entre lo que la justicia exige y lo que debe ser hecho, que las demandas de la moralidad tienen un estatus especial, que hay una conexión estrecha entre los resultados horribles de un acto y el que sea incorrecto, etc. ¿Qué otra cosa puede decirse a tal persona más que ella «simplemente no entiende?» (Jackson, 2012:83).

15 También aquí el naturalismo recurre al equilibrio reflexivo, pero en este caso en la parte analítica. Por medio del equilibrio reflexivo se perseguiría pasar de la moralidad *folk* a una moralidad *folk* madura (o varias, si es que no se converge en una).

vez que establecemos todas las conexiones que tiene que cumplir «bueno», buscaremos determinar qué propiedad natural, colocada en lugar de «bueno» puede cumplir ese rol.

En conclusión, el naturalismo analítico no establece a priori a qué propiedad natural hace referencia un término ético. El análisis a priori establece todas las conexiones de un término ético con una red de otros términos. Una vez que sabemos qué relación tiene un término como *correcto* con otros términos como *bueno*, *a posteriori* deberemos buscar qué propiedad natural cumple con el rol que cumple *correcto* en esa red.<sup>16</sup>

El proyecto es, entonces, sistematizar la moralidad *folk* para luego pasar a un estudio empírico.

### Naturalismo metodológico

El naturalismo metodológico, como destaqué al comienzo, rechaza una distinción tajante entre ciencia y filosofía. Así, rechazará que haya un método propio para la ética, que la vuelve ajena a la ciencia. Entre esos métodos, puede pensarse en el recurso a algún tipo de intuición, el análisis conceptual y el equilibrio reflexivo, entendido como un método para volver coherentes nuestros principios y juicios, más allá de los hechos.

En qué sentido se aplica una misma metodología en ciencia y filosofía ya resulta menos claro. En el caso de Sturgeon, que se considera un naturalista metodológico, además de sustantivo, vimos que adopta una concepción rea-

---

<sup>16</sup> Los naturalistas sintéticos consideran que están a salvo de la crítica de Moore porque ellos no sostienen que la identidad entre términos éticos y términos naturales se establece a priori. Jackson, en cambio, sí considera que esa identidad se establece a priori. La respuesta de Jackson a la objeción de Moore es la siguiente: «Es cierto, de acuerdo al funcionalismo moral, una historia descriptiva suficientemente rica lleva a priori a que un acto sea correcto; pero esto será un claro caso de una implicación a priori, conceptual, no obvia, precisamente debido a la complejidad de la historia moral funcionalista. Del mismo modo en que podemos dudar razonablemente del resultado de una suma numérica larga y compleja debido a que tiene sentido dudar de que la suma haya sido hecha correctamente (y, consecuentemente, insistir en que una afirmación acerca del resultado de la suma es una posición «sustantiva» en aritmética, y que está «abierta» a que su respuesta sea discutida), del mismo modo es razonable dudar del resultado de la compleja historia que el funcionalismo moral dice que lleva desde lo descriptivo hacia lo ético. Es más, lo que importa de acuerdo al funcionalismo moral es la naturaleza de la moralidad *folk madura*; la forma de la teoría en la que la moralidad *folk* actual se convertirá (o se convertiría), o en la que convergiría, bajo el efecto de la reflexión crítica. Puesto que eso es lo que establece el rol de la corrección, y, de este modo, una vez combinada con los hechos relevantes acerca de qué propiedades ocupan qué roles, el tipo de propiedad que la corrección es. Así, inevitablemente habrá, aquí y ahora, un grado sustancial de “apertura” inducido por el hecho mismo de que el rol de la corrección está actualmente bajo negociación» (1998:151).

lista del equilibrio reflexivo que sería de aplicación tanto en la ciencia, como en la ética.

En lo que sigue presentaré el modo en que Railton considera que la filosofía va «en tándem» con la ciencia, y terminaré con el desarrollo de una posición que da cuenta de que debe trabajarse «desde dentro» de la ciencia.

### El naturalismo sintético reduccionista: Peter Railton

En relación con Sturgeon, Railton invierte la relación entre el naturalismo metafísico y el naturalismo metodológico:

aunque, de hecho, el naturalismo contemporáneo del que nos ocuparemos aquí es tanto metodológico como sustantivo, debe lealtad primeramente al naturalismo metodológico. No es una visión impulsada por consideraciones metafísicas acerca de qué tipos de entidades existen o pueden existir, y por lo tanto no comienza con una definición de «propiedad natural» con el propósito de dividir el reino de lo posible en Real e Irreal. (Railton, 1989:156)

Su idea es lograr con los términos éticos lo que se logró con términos como «agua», para evitar que sean considerados como «flogisto». Cuando logró establecerse la identidad entre el agua y  $H_2O$ , se dio una definición reductiva de «agua», una definición que si bien modificó la definición anterior, reivindicó el uso del término. Por otro lado, los intentos infructuosos de reducción en el caso de términos como «poliagua» o «flogisto» llevaron a su abandono.

Si logramos, entonces, una definición reductiva, podremos reivindicar los términos éticos, a la vez que habremos ofrecido una defensa de los mismos contra los eliminativistas que los ven como análogos a «poliagua» y «flogisto».

Desde el naturalismo metodológico podrá defenderse un realismo ético en la medida en que la postulación de hechos morales cumpla una función explicativa de ciertos rasgos de nuestra experiencia. Los hechos postulados, para cumplir con estándares realistas, deberán satisfacer dos condiciones: a) independencia: su existencia y características tienen que ser independientes de nuestro pensamiento, y b) retroalimentación: debemos poder interactuar con ellos.

La postulación de un mundo externo tiene sentido de cara a explicar la coherencia, estabilidad e intersubjetividad de nuestra experiencia sensible. Pero, ¿qué función explicativa puede tener la postulación de hechos morales?

Railton ofrecerá su respuesta, pasando primero por el análisis de la valoración no-moral, de la noción de que algo es deseable, es bueno para una persona. ¿Podemos ofrecer una definición reductiva de estos conceptos, de ma-

nera de identificarlos con hechos independientes, con los que interactuamos y que permiten ofrecer explicaciones de otros hechos?

Railton considerará, en primer lugar, la noción de «interés subjetivo», de deseo. Que tengamos un interés subjetivo en  $x$  quiere decir que  $x$  tiene una valencia positiva para nosotros, que dispara en nosotros una actitud positiva, una inclinación favorable. En este sentido, el interés subjetivo funciona como una cualidad secundaria, análoga a la sensación gustativa de la dulzura: son experiencias que supervienen sobre las cualidades primarias del sujeto y del objeto: quien percibe está constituido de tal manera que este tipo de objetos, en este contexto, genera tal sensación. Estas cualidades primarias son la base de reducción de la cualidad secundaria, la base objetiva que se corresponde con, y ayuda a explicar, las experiencias subjetivas.

Sin embargo, esta noción de interés subjetivo no puede ofrecer una explicación de la valoración no-moral, en tanto no puede explicar la fuerza normativa de la misma, su deseabilidad. El interés subjetivo es compatible, por ejemplo, con la ignorancia del sujeto, de manera tal que puede otorgar una valencia positiva a algo que no es bueno para él, que no hubiese deseado, de haber contado con más información.

Railton introduce, entonces, la noción de interés subjetivo objetivado. Si llamamos  $A$  a un individuo cualquiera,  $A+$  será  $A$  con un conocimiento completo y claro de sí mismo y de su ambiente, y con una racionalidad instrumental sin defectos. Si le preguntásemos a  $A+$  qué desearía que su yo no idealizado  $A$  persiguiese, de estar él en el lugar de  $A$ , podemos suponer que la respuesta partiría de cierta base de reducción, la constituida por los hechos sobre  $A$  y sus circunstancias, que  $A+$  consideraría para determinar lo que desearía en caso de estar en los zapatos de  $A$ .

El ejemplo de Railton es el siguiente. Supongamos que Lonnie, en el marco de un viaje por el extranjero, se siente nostálgico y con cierto malestar estomacal. En tal situación, se le antoja calmar su estómago con una bebida familiar: un vaso de leche. Podemos decir, entonces, que Lonnie desea un vaso de leche, que ese es su interés subjetivo.

Lonnie+ sabe algo que Lonnie no sabe, que la causa de su malestar, además de la nostalgia, es una deshidratación que suelen padecer los turistas. Con ese cuadro, ingerir un vaso de leche generaría un mayor malestar estomacal y aumentaría su deshidratación. Lonnie+ sabe, además, que una dieta líquida mejoraría la condición física de Lonnie e, indirectamente, reduciría su nostalgia. De manera que, de estar él en los zapatos de Lonnie, desearía que Lonnie adoptase una dieta líquida.

La base de reducción del interés de Lonnie+ está constituida por las circunstancias y constitución de Lonnie. Podemos considerar, dirá Railton, que

este conjunto de cualidades primarias es la base de reducción del interés objetivo de Lonnie.

De aquí se seguiría que «x es bueno para A (en un sentido no-moral) si y solo si x puede satisfacer un interés objetivo de A». Que sea verdadero o no que x pueda satisfacer un interés objetivo de A dependerá de las bases de reducción, esto es, de la constitución de A y sus circunstancias. Podemos ver, entonces, que el hecho de que algo sea bueno para A no depende de que A piense que así lo sea: una dieta líquida era buena para Lonnie, pero Lonnie no pensaba eso. De manera que esta propiedad cumpliría con el primer requisito antes destacado, el de independencia: algo es bueno para A, independientemente de que A lo considere así.

Esta definición de «bueno», ¿puede contribuir a alguna explicación? Railton compara la situación de Lonnie con la de Tad, quien en la misma situación que Lonnie adquiere un deseo de iniciar una dieta líquida y procede a hacerlo. A diferencia de Lonnie, veremos a Tad reanimado, y podremos explicarlo por el hecho de que, si bien ambos actuaron guiados por sus deseos, los de Tad fueron congruentes con su interés objetivo (quizá por casualidad, quizá porque sabía lo que era bueno para él), mientras que los de Lonnie no.

La congruencia entre deseos e intereses objetivos generalmente ayudará a explicar el grado de satisfacción que uno alcanza en su vida. Y aquí no importa si nosotros pensamos que nuestros deseos reflejan nuestro interés objetivo, sino si realmente lo reflejan. Nuevamente, se destaca así la independencia del interés objetivo, con respecto a nuestros pensamientos.

Pero también podemos ver cómo se da la segunda condición para poder hablar de hechos reales: la de retroalimentación. Supongamos que Lonnie tomó leche y pasó a sentirse peor; por casualidad se encuentra con un negocio que vende 7-up, una bebida que también lo conecta con su entorno familiar, y decide comprar una. Lo alcanza una sensación agradable al tomarla, por lo que decide comprar otras más. A la mañana siguiente observa que se siente mucho mejor, razón por la cual Lonnie puede asociar la superación del tipo de malestar que tenía con el tipo de bebida que es una 7-up. De esta manera, los deseos de Lonnie han evolucionado, de cara a volverse más congruentes con sus intereses objetivos, esto es, con lo que es bueno para él.

Este mecanismo que vincula nuestros deseos con nuestros intereses objetivos posee, ciertamente, como destaca Railton, un fuerte sustento teórico. Es claro que perseguimos nuestro bienestar, que nuestros deseos contribuyen a su logro, que pueden estar equivocados y que pueden corregirse en virtud de nuestra experiencia o de la adquisición de conocimientos. Railton concluye:

Si se postula un mecanismo deseos/intereses y si lo que es bueno para alguien (desde un punto de vista no-moral) se determina en función de cuál es su interés objetivo, entonces podemos decir que el valor objetivo es capaz de jugar un rol en la explicación del valor subjetivo, del tipo que necesita quien se considera realista naturalista en relación con el valor. (2003:16)

Hasta aquí hemos podido ver que la noción de valor objetivo, de bueno, permite explicar muchas cosas en relación con el valor subjetivo, lo efectivamente deseado. A su vez, el valor objetivo tiene bases de reducción. ¿Cómo encaja en todo esto el realismo moral? ¿Podemos ir más allá, dentro del naturalismo, de lo que es valor no-moral?

Railton considera que sí, y lo hará dando cuenta de las normas morales que establecen cómo deben ser las cosas, antes que del valor moral, en tanto entiende que intuitivamente resulta más implausible aceptar un naturalismo en torno a lo correcto, que en torno a lo valioso. ¿En qué puede consistir una explicación naturalista de cómo deben ser las cosas?

Su punto de partida serán las explicaciones, nuevamente fuera del terreno moral, que implican normas naturalizadas. Supongamos que alguien pregunte por qué colapsó el techo de la casa. Una respuesta puede ser: «Para una casa que tiene que soportar esta cantidad de nieve, las vigas debieron haber sido de, al menos, 2x8 y no de 2x6». La explicación es aceptable si suponemos una meta y la existencia de una norma que refiere a un proceso que permite alcanzar la meta. Este «debe» es hipotético y la expresión que lo contiene puede traducirse en términos naturales: si un techo tiene que aguantar (cierta cantidad de) nieve, tiene que tener vigas de, por lo menos, 2x8. Para explicar la caída del techo puedo apelar a este criterio (a su incumplimiento), criterio que, aun cuando se presente en términos naturales, tendrá un rol regulativo en el contexto de la arquitectura.

Lo mismo podemos ver, dirá Railton, en torno a las normas de racionalidad individual, y esto nos va acercando a la moralidad. Una persona tendrá ciertas metas y creencias sobre cómo alcanzarlas, y una persona racional será aquella que actúa de manera eficiente, dados esos deseos y esas creencias. Supongamos que una persona deseaba evitar un conflicto, creía que la manera de evitarlo era yéndose de la reunión y que se fue de la reunión. ¿Cómo podemos explicar su conducta, esto es, su retiro de la reunión? Apelando a su racionalidad, a su disposición a cumplir con las normas que le permiten lograr de manera eficiente sus metas.

Los diferentes grados de racionalidad (en virtud de la capacidad para calcular probabilidades o hacer inferencias, por ejemplo) también permitirán explicar por qué algunos obtienen más éxito que otros en el logro de sus metas.

Por otra parte, nuestra irracionalidad tendrá efectos indeseados por nosotros, mientras lo contrario ocurrirá con nuestra racionalidad. De aquí que podamos llegar a adquirir creencias y estrategias que son racionales, sin que hayamos tenido que desplegar capacidades deliberativas. Podemos adoptar heurísticas inconscientes que son racionales, pero que no las adoptamos por un proceso deliberativo consciente.

También podemos apelar a la racionalidad individual para explicar conductas, pero ya no desde las metas y creencias que efectivamente tenga la persona, sino desde sus intereses objetivos.

Podemos hablar, entonces, de un realismo normativo:

Las evaluaciones de grados de racionalidad instrumental juegan un rol destacado en nuestras explicaciones del comportamiento individual, pero simultáneamente tienen fuerza normativa para el agente. Más allá de otras cuestiones que el agente pueda considerar, seguramente cuenta para él como un rasgo positivo de una acción el que sea eficiente en relación con sus creencias y deseos o, en el sentido extendido, eficiente en relación con creencias y deseos que reflejen apropiadamente su condición y circunstancias. (Railton, 2003:20)

Así como un arquitecto tiene sus metas, como individuos tenemos las nuestras y hay criterios que cumplir para su logro. Y nuestro sistema motivacional está diseñado para perseguir ese logro. De manera que las normas de racionalidad son tanto explicativas (de logros y fracasos), como normativas.

Dicho esto, podemos sí abordar las normas morales.

Railton primero presentará lo que considera el carácter distintivo de las normas morales, en relación con otros criterios de evaluación. En primer lugar, lo que se evalúa son las conductas o caracteres en situaciones en las que los intereses de más de una persona están en juego. En segundo lugar, la evaluación se hace desde un punto de vista imparcial, con igual consideración por los intereses de todos los potencialmente afectados. En este contexto, las normas morales reflejarían una racionalidad desde el punto de vista social.

Railton introduce, entonces, su criterio de corrección moral: lo que sería racionalmente aprobado, considerando por igual los intereses de todos los individuos potencialmente afectados, con información completa y clara.

Presentado el criterio, queda analizar su capacidad explicativa y la satisfacción de las dos condiciones para considerarlo parte de la realidad.

En primer lugar, si no tenemos en cuenta la corrección moral, habrá consecuencias, del mismo modo en que el hecho de que Lonnie no tenga en cuenta su necesidad de una dieta líquida (en virtud de que no sabe lo que es bueno para él) tiene consecuencias. En este caso, una manera de no tener en

cuenta la corrección moral es fomentar instituciones sociales que no tienen en cuenta los intereses de algún grupo. La consecuencia de desestimar los intereses de un grupo es el establecimiento de un potencial malestar social que, en ciertas circunstancias, se volverá manifiesto.

La existencia de instituciones sociales moralmente incorrectas no depende de que las personas entiendan que efectivamente son moralmente incorrectas. Puede darse el caso de que todos crean que las instituciones son justas, pero en la medida en que no tienen en cuenta los intereses de un sector, cabe esperar encontrar en los miembros de este sector un potencial para la insatisfacción (por ejemplo, puede considerarse que están alienados, lo cual, en ciertas circunstancias, dará lugar a la manifestación del descontento) que puede llevar a que, en algún momento, cambien sus creencias en torno a la justicia de esas instituciones.

La noción de racionalidad social juega un rol explicativo en tanto el alejamiento de la misma permite explicar ciertos fenómenos.

A su vez, «El descontento generado por el incumplimiento de la racionalidad social puede producir una retroalimentación que, a nivel social, promueve el desarrollo de normas que se aproximan más a la racionalidad social» (Railton, 2003:24).

Esta teoría explicativa también permite predecir. Ciertas circunstancias, más que otras, darán lugar a presiones que lleven a prácticas sociales que incorporen la consideración de intereses hasta ese momento no tenidos en cuenta:

Dada una caracterización adecuada de las condiciones que prevalecieron durante los procesos de evolución normativa descritos (...), la presente teoría afirma no sólo que esos cambios pudieron haber sido esperados, sino que una parte esencial de la explicación de que ocurriesen es un mecanismo por el cual los individuos que han visto negados sus intereses son empujados a formar valores comunes y a hacer causa común en función de intereses compartidos, presionando sobre prácticas sociales para aproximarse de manera más estrecha a la racionalidad social. (Railton, 2003:28)

Si bien esta teoría no ofrece imperativos categóricos, independientes de los deseos contingentes del agente, puede afirmar que debemos hacer algo, incluso cuando ello no ayude a satisfacer nuestros actuales fines o intereses objetivos. La determinación de lo que es moralmente correcto que haga, surge de lo que es racional desde un punto de vista social, no personal. Sin embargo, como con la aplicación de la lógica, dada la relevancia del hecho de que nuestra conducta pueda ser justificada desde un punto de vista del interés general, no resulta arbitrario asignar importancia a la moralidad.



Railton ofrece, entonces, definiciones reductivas y reivindicativas de los términos morales, que pueden introducirse en explicaciones de hechos naturales. Esta manera de trabajar no diferiría de lo realizado en otras ciencias y, si el trabajo realizado es correcto, ofrece una defensa del naturalismo sustantivo.

## El naturalismo evolutivo

Este enfoque asume la defensa de un naturalismo metodológico, en tanto los problemas filosóficos se abordan «desde dentro» de la ciencia. No hay otra metodología para la adquisición de conocimiento que la aplicada en ciencia. Y las hipótesis sobre lo que aún no sabemos deben surgir del conocimiento acumulado en las ciencias.

Las posiciones naturalistas desarrolladas en la segunda mitad del siglo pasado, que acabamos de presentar, pueden verse como más preocupadas por ser compatibles con los conocimientos científicos, que por apoyarse en los mismos. De aquí que eviten apelar a entidades *sui generis*, o a nociones como las de *razón categórica* o *libertad*, extrañas a las ciencias naturales. Sin embargo, la consolidación de la teoría evolutiva y, adoptando este marco teórico, los relevantes desarrollos en genética, biología, primatología, psicología, neurociencia, entre otras ciencias, han llevado a muchas filósofas y filósofos a adentrarse en estas ciencias para ver qué consecuencias se siguen para la metaética y la ética.

Antes que ofrecer una definición estricta de la ética, como la presentada por Railton, o a partir de la moralidad *folk*, se intenta determinar cuál puede ser su origen y qué consecuencias se siguen de ello. De todos modos, este enfoque *desde dentro* de la teoría evolutiva no ha dado lugar a consensos. Así, nos podemos encontrar con posiciones relativistas, como la de Jesse Prinz (2015); posiciones escépticas, como la de Richard Joyce (2006); posiciones racionalistas, como la de Joshua Greene (2013), etc. En lo que sigue, ilustraré este enfoque con la defensa de una posición que le otorga un lugar central al no-cognitismo.

Tomaré como punto de partida las investigaciones del neurocientífico Antonio Damasio sobre la valoración. En *Self Comes to Mind* parte del hecho de que durante mucho tiempo los únicos seres vivos que habitaron el planeta fueron organismos carentes de conciencia, e incluso de toda célula nerviosa. Aquellos seres vivos debieron superar diferentes escollos para lograr sobrevivir y reproducirse, y lo lograron porque contaron con dispositivos que les permitieron permanecer dentro de un intervalo homeostático, esto es, que les permitieron mantener ciertos parámetros (de temperatura, energía, etc.) dentro

del rango requerido para mantenerse con vida, y salirse del cual implica perecer. A medida que los organismos se fueron volviendo más complejos, también se complejizaron los dispositivos que permitían mantener el equilibrio homeostático necesario para sobrevivir y reproducirse.

En este contexto podemos dar sentido a un término como «valor». En palabras de Damasio:

Para cualquier organismo considerado como un todo, lo que tiene un valor primordial es, dicho sin ambages, la supervivencia con una salud buena hasta una edad compatible con el éxito reproductivo. La selección natural ha perfeccionado la maquinaria de la homeostasis con el fin precisamente de permitir que así sea. En consecuencia, el estado fisiológico de los tejidos de un organismo vivo, en el interior de un intervalo homeostático óptimo, es el origen más profundo del valor biológico y las valoraciones. Y lo mismo se puede afirmar tanto de los organismos pluricelulares como de aquellos organismos cuyo «tejido» vivo se limita a una sola célula. (2010:45-46)

A partir de esta valoración primordial, el aprendizaje lleva a los organismos a asignar valor a los medios y obstáculos. Dicho de manera sencilla, esto quiere decir que los organismos se comportarán hacia los hechos asociados de manera similar a como se comportan hacia los estímulos incondicionados: perseguirán los medios y evitarán los obstáculos.

Ahora bien, entre los dispositivos que permiten mantener el equilibrio homeostático debemos colocar a la conciencia:

la conciencia surge gracias al valor biológico, como colaboradora en la gestión más efectiva del valor. Pero la conciencia no inventó el valor biológico ni el proceso de valoración. Con el paso del tiempo, en las mentes humanas, la conciencia reveló el valor biológico y permitió el desarrollo de nuevas maneras y medios de gestionarlo. (Damasio, 2010:31)

La conciencia habría sido seleccionada como una herramienta más para el logro de lo que es biológicamente valioso. ¿De qué manera podría ser útil la conciencia? Puesto que la conciencia no tiene un registro directo de lo que ocurre en el cuerpo, no puede ser la encargada de tareas esenciales como, por ejemplo, regular la liberación de hormonas o neurotransmisores para el mejor funcionamiento del cuerpo. Sin embargo, la corteza somatosensorial va generando «mapas» de lo que ocurre en el cuerpo, los cuales a su vez tienen por efecto la generación de sentimientos conscientes.

El correlato consciente de encontrarnos en un punto óptimo del intervalo homeostático es un sentimiento de bienestar. Así, cuando conscientemente estamos persiguiendo nuestro bienestar, o evitando todas las situaciones que nos generan un sentimiento de malestar, estamos persiguiendo los medios que contribuyen a ubicar a nuestro cuerpo en un intervalo homeostático óptimo, de cara a sobrevivir y reproducirnos.

A su vez, el aprendizaje por asociación cada vez resultará más complejo. Una vez que nuestra capacidad cognitiva establece que algo es un medio para el logro de lo que es considerado valioso, ese medio puede adquirir, a su vez, valor y oficiar de guía para un nuevo esfuerzo cognitivo.<sup>17</sup> Como destacan Patricia Churchland y Christopher Suhler:

al embellecer la antigua organización del sistema de recompensa subcortical con un sofisticado *input* cortical, un plan puede ser evaluado por sus probables consecuencias. Un *input* cortical más rico permite predicciones y evaluaciones más

---

17 En realidad, para que esto ocurra no hace falta contar con capacidades cognitivas complejas. Todo proceso de aprendizaje tendría este resultado. Read Montague (2006) ha abordado los procesos cerebrales por los que puede darse este traslado de valoración de fines a medios, a partir de las investigaciones con monos de Wolfram Schultz. En el trabajo de Schultz se observaba la actividad de ciertas neuronas dopaminérgicas ante la llegada de una recompensa (jugo), ante una anticipación de la recompensa (encendido de una luz) y ante un error en la predicción de la recompensa. Inicialmente, cuando la luz se presenta no hay cambios en la actividad de las neuronas dopaminérgicas, puesto que la luz no está asociada a ninguna recompensa; esto es, inicialmente, la luz es asociada con una señal «las cosas son tal como lo esperado», una modificación del ambiente que no implica mejoras ni castigos. Sin embargo, la llegada del jugo sí causa un estallido de actividad en las neuronas dopaminérgicas; esto es, el envío de jugo es inicialmente asociado con una señal «las cosas son mejores que lo esperado» en las neuronas. Si los pares luz-jugo son enviados repetidamente al sujeto, se dan dos cambios destacados. La respuesta inicial asociada al jugo (la respuesta «las cosas son mejor que lo esperado») desaparece literalmente y las neuronas ya no cambian su actividad cuando el jugo llega luego de la luz. Esto significa que el sistema dopaminérgico ha aprendido a esperar el tiempo y la cantidad de jugo enviado, y lo demuestra no cambiando su actividad («las cosas son tal como lo esperado») cuando llega el jugo. Pasemos al segundo cambio. La falta de respuesta inicial a la luz (la respuesta «las cosas son tal como lo esperado») cambia hacia un estallido de actividad. Las neuronas ahora informan que la luz es «mejor que lo esperado»; esto es, reaccionan a la luz de la misma manera en que, inicialmente, reaccionaban al jugo. En algún sentido, el valor del jugo ha sido transferido a la luz, la cual deviene un «valor proximal» para la llegada futura de jugo que predice. La respuesta neural a la luz se transforma en una especie de nota neural promisorio para el real valor (agua azucarada) que llegará pronto. Este traslado de la respuesta de estallido desde el jugo hacia la luz es comunicado por los axones diseminados de las neuronas dopaminérgicas, diciéndole al resto del cerebro que trate a la luz como si tuviese el mismo valor de recompensa intrínseca que el jugo. De todos modos, si luego deja de enviarse jugo tras prender la luz, habrá una reconfiguración neuronal que tendrá por resultado que la luz nuevamente carezca de valor. En el caso del *homo sapiens*, la dinámica es similar, aunque la complejidad aumenta.

ricas. Unas metas pueden anidarse dentro de otras. Los planes pueden volverse muy elaborados y las metas muy abstractas. Al hacer uso de patrones de causalidad aprendidos, el cerebro puede ensamblar evaluaciones de las consecuencias de un plan. Estas son modelos «si..., entonces», y pueden volverse realmente muy sofisticadas. (2014: 318-319)

En general, este mecanismo es lo suficientemente flexible como para poder aprender de los diversos entornos. El fin último es estable, pero todo lo que adquiera valor por su relación con él dependerá de las circunstancias en las que se dé el desarrollo del organismo. Todo esto da lugar a que, incluso si en última instancia todos valoramos lo mismo, lo cierto es que la experiencia y la cultura en la que nos desarrollamos nos llevarán a asignarle valor a distintos medios, y dado el anidamiento de metas que a nivel cognitivo podemos generar, muchas veces se darán conflictos muy complejos entre personas, ya que cada una puede estar defendiendo un juicio al que asocia fuertes sentimientos, que considera tan final como el disparado por nuestra búsqueda de supervivencia.

Concentrémonos, ahora, en el origen de los valores morales. En la misma línea de Damasio, Churchland destaca que

Todos los sistemas nerviosos se organizan para cuidar de la supervivencia básica del cuerpo del que forman parte. Desde una perspectiva evolutiva, el punto general es claro: se selecciona el «cuidado» de uno mismo por sobre el «descuido» de uno mismo. (2011:27)

Obviamente, los organismos que no han sido capaces de cuidar de sí mismos no están hoy entre nosotros. Las especies que han llegado hasta nuestros días son las que han logrado cuidar de sí mismas.

Pero para lo que nos interesa, lo crucial es el siguiente paso:

en el caso de los mamíferos, los ajustes evolutivos en los sistemas emocionales, endocrino, de estrés y de recompensa/castigo efectivamente extienden el rango de individuos por los que el animal se preocupa (...) la madre rata se comporta como si sus recién nacidos estuvieran incluidos en su ámbito básico homeostático. (Churchland, 2011:30)

Este paso es determinante. Ahora, el bienestar de un animal depende del bienestar de otro. En alguna medida, dos animales se han vuelto una unidad. Ya no solo tendrán valor positivo las cosas y circunstancias que contribuyen a la supervivencia y reproducción del organismo que valora, sino que también

tendrán valor positivo las cosas y circunstancias que afecten favorablemente a otro organismo y a la relación entre ellos.

Una vez que comprendemos este pasaje de la preocupación por el propio organismo a la preocupación por otro, podemos hacernos una idea de cómo los lazos pueden extenderse:

la modificación que nos lleva a cuidar de otros que son nuestra descendencia, puede experimentar modificaciones adicionales, tal vez de carácter menor, hasta extenderse al cuidado de individuos que no son de nuestra propia descendencia pero cuyo bienestar es de alguna manera significativo para el bienestar propio y el de nuestra descendencia. (Churchland, 2011:32)

No debe pensarse que el cuidado es estratégico, esto es, que cuidamos a otro porque ello contribuirá a nuestra supervivencia y bienestar. El camino es inverso. Se generan mutaciones que hacen que nuestro bienestar esté asociado al bienestar de otros, razón por la cual los ayudamos, y una consecuencia de este hecho es que generamos comportamientos cooperativos mutuamente beneficiosos.

Las relaciones de cooperación y confianza pudieron extenderse y, en el caso de los humanos, dar lugar a pequeñas tribus de cazadores-recolectores. En el marco de estas comunidades podían tener completo sentido las pautas normativas sobre qué debe hacerse y qué no. Al existir una valoración de las relaciones y el bienestar de cada uno, volcarán su capacidad cognitiva a establecer los mejores medios para alcanzar el bienestar de todos. A su vez, esto puede dar lugar a debates en torno a cuáles son los mejores medios, lo que derivará en debates sobre deberes, sobre normas morales y, en última instancia, sobre cuáles pueden ser los mejores desarrollos culturales que faciliten el bienestar general.<sup>18</sup>

Estas investigaciones parecerían fortalecer una metaética naturalista de corte humeano-darwiniano, como la defendida por Oliver Curry (2006). De Hume se adopta la idea de que son nuestras pasiones, nuestros sentimientos, los que determinan lo que es valioso. Somos sujetos morales porque tenemos ciertas pasiones asociadas al bien común (nos sentimos bien ayudando, nos sentimos mal ante la mera imaginación de matar a alguien con nuestras manos, etc.), y no porque tengamos una razón que determina que es correcto ayudar o que está mal matar y que, de alguna manera, nos motiva a actuar.

---

18 De todos modos, detrás de las enormes diferencias culturales que pueden darse, existe una biología prácticamente idéntica.

De Darwin se adopta la explicación de por qué tenemos esos sentimientos. Hemos llegado a tenerlos en virtud de la evolución por selección natural.

La evolución por selección natural ha dado lugar a un animal con ciertos sentimientos que, luego, son la base sobre la que podemos llegar a construir una teoría moral. En alguna medida deseamos el bien común, lo que explica que podamos adherir a una teoría ética que exija los comportamientos que contribuyan al bien común. Pero si no deseásemos el bien común, si fuésemos seres a quienes no les afecta en lo más mínimo el sufrimiento ajeno, ningún argumento racional nos llevaría a adherir a normas que exigiesen que nos preocupemos por quienes sufren.

Desde esta perspectiva, lo que efectivamente es valioso será aquello que efectivamente sintamos de esa manera. Luego, buscaremos qué medios contribuyen a lograr aquello que consideramos valioso, y una vez determinados los medios sabremos lo que debemos hacer.

En un argumento moral encontraríamos estos elementos: una premisa que afirma qué es valioso para nosotros (lo cual viene determinado por nuestra naturaleza emocional), una segunda premisa que determina cuál es el medio para lograr ese fin valioso (medio que está determinado por cómo es el mundo y que conoceremos gracias a nuestra capacidad cognitiva, que nos permite conocer qué es causa de qué) y una conclusión que se enuncia como un deber.<sup>19</sup>

Este naturalismo entiende que una teoría ética debería establecer, en primer lugar, qué es lo que realmente valoramos indagando en nuestra naturaleza emocional (la cual, para poder hablar de ética, tendrá que incluir emociones que promuevan el bien común). En segundo lugar, la teoría ética debería hacer uso del conocimiento científico relevante para establecer cuáles son los mejores medios para lograr lo que realmente valoramos. Y, en tercer lugar, y teniendo en cuenta lo que valoramos y los medios idóneos para lograrlo, podrá concluir cuáles son nuestros deberes.<sup>20</sup>

Pero esto nos lleva a un punto central: ¿qué fundamenta a los valores? Este naturalismo ofrece una explicación evolutiva del origen de los valores. Es azaroso que un mamífero valore una cosa y otro mamífero valore otra cosa. Un mamífero puede valorar las frutas y otro mamífero valorar la carne de otro animal. Todo esto es el producto de la mutación genética en su interacción con el medioambiente (que incluye, obviamente, a otros seres vivos). Este es el origen de los valores. Luego, las capacidades de cada animal le ayudarán a encontrar medios adecuados para conseguir lo que considera valioso.

---

19 No tiene por qué ser una única premisa en cada caso. Pueden ser conjuntos de premisas.

20 Para un mayor desarrollo de esta posición y de su aplicación a problemas éticos, ver Elgarte y Daguerre (2021).

Desde el enfoque humeano-darwiniano, no tiene sentido preguntarse si está justificado valorar lo que, en última instancia, valoramos. Sí tiene sentido preguntarse algo distinto, a saber, si está justificado legalizar el aborto o no (esto es, si debe legalizarse o no), si debemos combatir el patriarcado o no, si debemos redistribuir la riqueza o no. Los argumentos tendrían esta forma: dado que valoramos x, y puesto que redistribuir la riqueza contribuye (o no) a lograr x, debemos (o no) hacer x.

Pero, ¿cómo podemos justificar lo que consideramos valioso en última instancia? El humeano-darwiniano entiende que, en última instancia, diremos que valoramos algo porque sí, porque así lo sentimos. En relación con esta valoración, lo único que podemos hacer es ofrecer una explicación de por qué lo sentimos así, pero nunca una justificación.

¿Por qué el humeano-darwiniano tendría que buscar otra cosa que una explicación de los valores últimos? En todo caso, ¿no recae el peso de la prueba en quienes asumen la existencia de un mundo «ideal», o de un dios creador de la moral, o de un sujeto racional libre que establece las reglas de convivencia sin estar determinado por el mundo natural?

Desde este enfoque, todos los juicios éticos sinceros expresarían un sentimiento positivo o negativo. Los juicios sobre valores últimos solo expresarían sentimientos. De aquí que podamos ver en estas investigaciones un apoyo al no-cognitivismo. Sin embargo, todos los juicios sobre medios no solo arrastran el sentimiento positivo asociado al fin, sino que expresan una creencia que puede ser verdadera o falsa. En relación con los medios, puede hablarse de conocimiento ético, de expresiones que apuntan a reflejar lo que ocurre en el mundo, que pueden testearse como cualquier afirmación científica.

Hasta aquí, entonces, un ejemplo de cómo hacer metaética partiendo de investigaciones en ciencias naturales (siempre en sentido amplio). Como destacó anteriormente, esto no quiere decir que haya un consenso en cuanto a las conclusiones. El consenso es, más bien, metodológico: se parte de los resultados de las investigaciones científicas y se sacan conclusiones, con un mayor o menor grado de especulación, sobre cuestiones metaéticas.

## **Conclusiones**

En resumen, el naturalismo ético puede entenderse de diversas maneras. En todos los casos se supone un vínculo entre la ética y las ciencias naturales, tomadas en sentido amplio.

Dando por supuesto que las propiedades éticas son propiedades naturales, podemos tratarlas como a otras propiedades naturales que no somos capaces de

reducir. En vez de determinar de manera directa qué las constituye, intentaremos establecer sus conexiones causales con otras propiedades naturales, lo cual no solo permitirá mejorar las explicaciones científicas, sino que también podrá arrojar claridad, con el tiempo, sobre qué constituye a esas propiedades éticas.

Otro modo de proceder es suponiendo que el significado de los términos éticos viene dado por sus conexiones con otros términos éticos y no-éticos. Corresponde, entonces, que hagamos un análisis de los significados de los términos éticos, que establezcamos su identidad con términos descriptivos, no-éticos, y que recién ahí determinemos empíricamente qué propiedad natural cumple con tener las conexiones que definen a cada término ético.

En el primer caso estaríamos viendo a la ética como una ciencia natural más. En el segundo caso, la ciencia natural tendría una aparición posterior al desarrollo del análisis ético propiamente filosófico.

Si en vez de dar por supuesto que las propiedades éticas son propiedades naturales, intentamos establecerlo mediante el método de las ciencias naturales, podemos considerar que lo que corresponde hacer es ofrecer definiciones reductivas de los términos éticos, que sean reivindicativos de los mismos, así como el descubrimiento de la identidad de *agua* y  $H_2O$  fue reivindicativa del uso del término *agua*. La ética, en este sentido, iría en tándem con las ciencias naturales, lo cual querría decir que es una réplica de las ciencias más básicas, en un lenguaje más abstracto.

Una opción diferente es entender a la ética desde dentro de las ciencias naturales. La teoría científica actualmente más relevante para este fin sería la teoría de la evolución. El objetivo es comprender las nociones de *bueno* y *correcto*, por ejemplo, en el marco de la evolución por selección natural. Previo a la eclosión de desarrollos biológicos, psicológicos, neurocientíficos, etc. que han tomado como marco la teoría de la evolución, una eclosión que tuvo lugar en las últimas décadas, los caminos plausibles podían consistir en establecer conexiones causales, arriesgar definiciones reductivas o recurrir a un análisis de la moralidad *folk*. No obstante, el ingente desarrollo de los conocimientos científicos sobre la naturaleza humana, ha llevado a muchos de quienes han escrito sobre estos temas en el siglo XXI a considerar más promisoría la adopción de un enfoque como el que describimos en último término, cuyas raíces se hunden mucho más directamente en el creciente *corpus* de investigaciones científicas sobre la naturaleza humana.



## Referencias bibliográficas

- Boyd, R.N. (1988). How to Be a Moral Realist. En Sayre-McCord, G. (Ed.), *Essays on Moral Realism* (pp. 181–228). Cornell University Press.
- Churchland, P. (2011). *Braintrust*. Princeton University Press.
- Churchland, P. y Sulher, Ch. (2014). Agency and Control: The Subcortical Role in Good Decisions. En Sinnott-Armstrong, W. (Ed.), *Moral Psychology. Vol. 4: Free Will and Moral Responsibility* (pp. 309–326). MIT Press.
- Curry, O. (2006). Who's Afraid of the Naturalistic Fallacy? *Evolutionary Psychology*, 4, 234–247.
- Damasio, A. (2010). *Self Comes to Mind*. Pantheon Books.
- Elgarte, J. y Daguerre, M. (2021). Construyendo desde adentro: repensando la metaética y el debate sobre el aborto desde una comprensión evolutiva de la naturaleza humana. En López Orellana, R. y Suárez Ruiz, E.J. (Eds.), *Filosofía posdarwiniana* (pp. 349–374). College Publications.
- Greene, J. (2013). *Moral Tribes*. The Penguin Press.
- Harman, G. (1977). *The Nature of Morality*. Oxford University Press.
- Harman, G. (1986). Moral Explanations of Natural Facts –Can Moral Claims Be Tested Against Moral Reality? *The Southern Journal of Philosophy*, XXIV, 57–68.
- Jackson, F. (1998). *From Metaphysics to Ethics*. Oxford University Press.
- Jackson, F. (2005). What are cognitivists doing when they do normative ethics? *Philosophical Issues*, 15, 94–106.
- Jackson, F. (2012). On ethical naturalism and the philosophy of language. En Nuccetelli, S. y Seay, G. (Eds.), *Ethical Naturalism* (pp. 70–88). Cambridge University Press.
- Joyce, R. (2006). *The Evolution of Morality*. The MIT Press.
- Montague, R. (2006). *Why chooses this book?* The Penguin Press.
- Moore, G.E. (1903). *Principia Ethica*. Cambridge University Press.
- Prinz, J. (2015). Naturalizing Metaethics. En Metzinger, Th. y Windt, J. (Eds.), *Open Mind* (pp. 1146–1172). Frankfurt am Main: MIND Group.
- Railton, P. (1989). Naturalism and Prescriptivity. *Social Philosophy & Policy*, 7(1), 151–174.
- Railton, P. (2003). *Facts, Values, and Norms*. Cambridge University Press.
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.
- Sturgeon, N.L. (1986). Harman on Moral Explanations of Natural Facts. *The Southern Journal of Philosophy*, XXIV, 69–78.
- Sturgeon, N.L. (2006a). Ethical Naturalism. En Copp, D., *The Oxford Handbook of Ethical Theory* (pp. 91–121). Oxford University Press.
- Sturgeon, N.L. (2006b). Moral Explanations Defended. En Dreier, J. (Ed.), *Contemporary Debates in Moral Theory* (pp. 241–262). Blackwell Publishing Ltd.

## 2. Constructivismo moral

*Martín Oliveira*

Este capítulo está dedicado a reconstruir la posición que se conoce, en el ámbito de la filosofía práctica, como «constructivismo moral» o «constructivismo ético». A esos efectos, la segunda sección del texto describe el surgimiento de esta posición en la literatura especializada, su tesis básica y su lugar entre las posiciones clásicas de la metaética. La tercera sección desarrolla una taxonomía para clasificar las distintas variantes del constructivismo, así como algunos de los ejemplos más importantes de ellas. En la cuarta sección se presenta una importante objeción en forma de dilema que se suele elevar frente al constructivismo, las implicaciones que ella tiene y la mejor estrategia de respuesta a este problema, el constitutivismo. Finalmente, la conclusión sugiere una prospectiva para el constructivismo moral, sobre todo cuando se lo concibe desde una perspectiva «práctica» en filosofía moral.

### El origen y la tesis básica del constructivismo

La literatura especializada suele atribuir a John Rawls la primera aplicación del término «constructivismo» dentro de la disciplina de la filosofía práctica (Darwall, Gibbard y Railton, 1992:138; Bagnoli, 2013a:1). En el extenso artículo *Kantian Constructivism in Moral Theory*, el filósofo norteamericano desarrolló un aspecto que, de acuerdo con su juicio, no había enfatizado en su conocida obra *A Theory of Justice*. Este aspecto no era otro que una particular concepción de la objetividad de los juicios morales, específicamente pensada para salvar el problema del desacuerdo entre teorías morales sustantivas.<sup>1</sup> En rigor, Rawls afirmó que: «El constructivismo kantiano sostiene que la objetividad moral debe ser entendida en los términos de un punto de vista social que todos puedan aceptar. Más allá del procedimiento para construir principios de justicia no hay hechos morales» (1980:519).

---

1 Es necesario aclarar que Rawls volvería a distinguir su propia concepción de la objetividad de los juicios morales incluso de la posición que en este artículo denominó «constructivismo kantiano». Específicamente, denominó a su particular concepción de la objetividad, adecuada para el dominio de lo político, «constructivismo político» frente al más amplio constructivismo «moral» kantiano. Cfr. Rawls, 2005:90.

De todos modos, Ronald Dworkin en su artículo *Justice and Rights* ya había sugerido la lectura del argumento de la «posición original» de Rawls como un «modelo constructivo» (Dworkin, 1973:160).

En buena medida, la definición del constructivismo que produjo Rawls resulta muy útil para distinguir las distintas tesis y formas en las que se reconstruyó el constructivismo desde entonces. Por un lado, como por ejemplo se evidencia en la primera oración antes citada, ha sido entendido como una tesis acerca de lo que hace verdaderos o correctos los juicios morales.<sup>2</sup> Por el otro, la segunda oración de la definición rawlsiana sugiere que puede entenderse al constructivismo como una posición normativa acerca de la clase de juicios morales que debemos aceptar en ciertos campos prácticos en particular.<sup>3</sup>

La primera forma de entender al constructivismo lo describe claramente como una posición metaética que afirma una tesis específica sobre el tipo de elemento que sirve de razón o fundamento para la verdad o corrección de nuestros juicios morales. Específicamente, el constructivismo metaético sostiene que la verdad o corrección de los juicios morales depende de que estos sean producto de un procedimiento de construcción o, más en general, del ejercicio del «punto de vista práctico», definido por ejemplo como el uso autónomo de la razón práctica o el resultado de un procedimiento de deliberación adecuado. El que los juicios morales verdaderos o correctos sean resultado de ese «punto de vista práctico» o procedimiento de construcción del que participan los agentes morales, a su vez, sería lo que los hace objetivos y autoritativos frente a ellos.<sup>4</sup> En este sentido, el constructivismo ha sido usualmente presentado como una tercera posición intermedia entre el realismo y el antirealismo morales. Su principal atractivo consistiría en retener cierta concepción de la objetividad de los juicios morales, que el antirealismo tiene severos problemas en explicar, y a la vez evitar problemas ontológicos y motivacionales, de los que es presa el realismo.

---

2 Si bien puede hacerse una distinción entre la verdad, corrección o validez de los juicios morales, en este trabajo los tomaremos como términos indistintos.

3 En menor medida ha sido entendido como una tesis epistemológica acerca del modo en el que se justifican creencias morales. Esta es la particular posición de Carlos Nino, quien a su vez la combina con una forma de constructivismo moral basada en ciertos rasgos formales de la «práctica del discurso moral». Cfr. Nino, 1988; 2007:92-128.

4 Como se desprende de esta caracterización, el constructivismo en su versión metaética ha sido entendido a partir de su respuesta a la pregunta por el tipo de elemento que puede servir de razón o fundamento a nuestros juicios morales y por ofrecer una caracterización de la objetividad de esos juicios a partir del procedimiento de construcción. El otro problema tradicional que normalmente se atribuye a la metaética, el del significado de los juicios o términos morales ha recibido menos atención desde las posiciones constructivistas. Con todo, Sharon Street ha sugerido que el constructivismo metaético sería compatible con una semántica inferencialista. De acuerdo con esta filósofa, el significado de los juicios morales se explicaría como «los tipos de inferencias que uno debe estar haciendo para contar como *[un agente]* que utiliza conceptos normativos» (Street, 2010:377; Bagnoli, 2013a:14).

La segunda manera de entender al constructivismo, es decir su variante normativa, es notablemente más modesta. En la medida en la que solo afirma una proposición normativa acerca de los juicios morales que debemos aceptar, a veces solo para un determinado campo práctico (la moral de lo correcto o lo incorrecto, o la justicia social, para mencionar algunos), evita pronunciarse respecto de las más tradicionales y controvertidas preguntas ontológicas y semánticas que pretende responder la metaética.

Se suele defender la plausibilidad del constructivismo a partir de dos argumentos negativos, fundamentalmente identificados como problemas cruciales del realismo moral. El primero de ellos se formula como un problema de carácter ontológico y epistemológico derivado de suponer que hay hechos en el mundo que hacen verdaderos o correctos nuestros juicios morales. Se trata del argumento de la «rareza» (*queerness*) de los hechos morales que John Mackie construyó como objeción al realismo intuicionista y no-naturalista de George Edward Moore. Para ser breves, la objeción afirma que el realista no-naturalista debe postular una ontología especial para explicar los hechos morales, si no se reducen a hechos naturales del mundo, y a su vez una facultad moral especial para aprehenderlos o conocerlos. Dichos postulados son extraños a una visión naturalista y científica del mundo (Mackie, 1977:38-40; Enoch, 2009:324; Bagnoli, 2013b:157). El segundo argumento es, si se quiere, una continuación o derivación del anterior. En particular, este argumento objeta que el realismo, comprometido con la tesis de que hay hechos morales independientes de los juicios de los agentes, no puede explicar el carácter intrínsecamente motivacional de los juicios morales. Esto sería consecuencia de que para el realista los juicios morales son eminentemente cognitivos y buscan representar con verdad esos hechos morales del mundo. Sin embargo, esta forma de entender los juicios morales no explica de ningún modo cómo motivan a los agentes morales<sup>5</sup> (Korsgaard, 1996:16; 2008:315). Ambas objeciones parecen conducir a una preocupación más general respecto al modo de entender el carácter y el rol de la filosofía práctica desde una perspectiva realista. Según ella, como se estableció antes, los juicios morales buscan representar hechos del mundo y su correspondencia con esos hechos los hace verdaderos y objetivos. En ese sentido los juicios morales tienen un contenido y función cognitivos que luego debemos aplicar en la práctica. Sin embargo, este modo teórico de entender la moralidad confunde conocimiento y acción y, más aún, de acuerdo a Korsgaard ya «presupone la habilidad de ac-

---

5 Se podría llegar a señalar que el argumento de la rareza de Mackie es especialmente fuerte contra las versiones no-naturalistas del realismo y no, en cambio, contra las versiones naturalistas del mismo. De todos modos, el realismo naturalista seguiría siendo blanco de la objeción respecto a su incapacidad para explicar el carácter motivacional de los juicios morales.

tuar» (2008:315). Más bien, el modo adecuado de entender la filosofía moral y política es atribuirle un carácter eminentemente práctico y de resolución de problemas (Rawls, 2005; Korsgaard, 2008:315). Es esta última manera de entender la filosofía moral la que el constructivismo pretende encarnar en su mejor interpretación o defensa.

Sin embargo, la mayoría de las defensas del constructivismo también se encuentran fundadas en un argumento positivo. Dicho argumento sostiene que los elementos fundamentales de la agencia práctica, como por ejemplo los de la objetividad de las normas morales y su rol motivacional, solo pueden ser comprendidos si adoptamos un «punto de vista práctico». La explicación filosófica de ese «punto de vista práctico», como cabe suponer, consiste en alguna de las variantes del constructivismo. Quizás la versión más clásica de ellas sea la del constructivismo que cabe encontrar en la obra de Immanuel Kant. El filósofo de Königsberg argumenta que el principal problema de las teorías morales de su época, el racionalismo dogmático y el sentimentalismo, fallan como teorías de la razón práctica en la medida en la que son heterónomas y por ende tienen consecuencias escépticas. La causa de ello es que ambas teorías derivan la autoridad de fuentes externas a la facultad misma de la razón, aunque tienen explicaciones diferentes para este fenómeno.<sup>6</sup> Si queremos concebir la posibilidad de obligaciones morales que tengan una genuina «necesidad racional» y alcance universal, la única opción es considerar, de acuerdo con las interpretaciones constructivistas de Kant, que son construidas por la propia razón práctica en tanto autónoma y autolegisladora como condición de posibilidad de ellas. Esta intuición fundamental ha servido para producir distintas versiones de la posición constructivista y explicar las consecuencias morales del llamado «punto de vista práctico».

### **Algunas variantes contemporáneas del constructivismo moral**

Como se mencionó previamente existen distintas variantes contemporáneas del constructivismo en la literatura especializada. Más aún, tampoco hay un consenso absoluto respecto a la forma de caracterizar y distinguir las distintas variantes del constructivismo. Mientras que algunos autores sostienen que to-

---

<sup>6</sup> El sentimentalismo falla porque parece reducirse a una concepción instrumental de la racionalidad práctica en el mejor de los casos que, en el léxico kantiano, apenas puede producir imperativos hipotéticos que unan medios afines. Estos fines, a su vez, no serían otros que los deseos o sentimientos contingentes de los agentes morales. El racionalismo dogmático, por su parte, «ignora el rol de la razón en hacer valiosos los fines de la acción» y así falla en explicar la autoridad de la razón. Cfr. Kant, 2012:103–108, IV:441–44; Bagnoli, 2013a:5–6.

das las variantes del constructivismo son posiciones normativas o sustantivas en la forma de un «proceduralismo hipotético» (Darwall, Gibbard y Railton, 1992:14), otros tantos autores hacen esfuerzos por defender la especificidad de la versión metaética del constructivismo (Enoch, 2009; Street, 2010; Bagnoli, 2013). Por otro lado, de manera perpendicular a la distinción entre el constructivismo metaético y el normativo, también es posible esbozar una distinción entre un constructivismo global respecto a todo el dominio de lo práctico y un constructivismo restringido a solo una esfera de la agencia práctica. Esto permite enumerar al menos cuatro maneras distintas de entender al constructivismo, aunque no todas hayan sido defendidas claramente. Una de ellas sería la del constructivismo metaético global. Se trataría de una tesis metaética acerca de la verdad de los juicios morales de todo el campo práctico. Su variante restringida sería una tesis metaética respecto de una clase limitada de juicios morales. Por otro lado, el constructivismo normativo global sería representado por una tesis normativa acerca del tipo de juicios morales que debemos aceptar en todo el campo de la agencia práctica. Mientras que, finalmente, un constructivismo normativo restringido estaría constituido por una tesis normativa acerca del tipo de juicios que debemos aceptar para un campo específico de la moralidad.

Quizás el que defiende Christine Korsgaard sea el ejemplo más acabado de constructivismo metaético, global contemporáneo y de inspiración kantiana.<sup>7</sup> Incluso es plausible sostener que la estrategia argumentativa general de la filósofa a favor de su variante del constructivismo es un ejemplo de los «argumentos trascendentales» de Kant. El argumento de Korsgaard comienza asumiendo que todo agente moral es un agente «auto reflexivo» y capaz de tener actitudes valorativas. La capacidad de auto reflexión supone la capaci-

---

7 La caracterización del constructivismo de Korsgaard, sin embargo, no es sencilla. Por un lado, ella afirma que «hay respuestas a las preguntas morales *porque* hay procedimientos correctos para llegar a ellas» (Korsgaard, 1996:37) y caracteriza a esta posición como un «realismo procedimental». Esto abona la idea de que el suyo es un constructivismo metaético que aspira a aplicarse sobre el dominio completo de la moralidad. Sin embargo, en trabajos más recientes la filósofa ha adoptado la antes mencionada «concepción práctica de la filosofía moral» la cual parece suponer que la metaética no tiene ninguna implicación en el campo de lo normativo y que todos los argumentos que pretendan validar conclusiones normativas deben ser así también normativos (Korsgaard, 2008). Si esto es así, su posición podría caracterizarse como un constructivismo normativo de alcance global, de hecho, así parece entenderla David Enoch. Cfr. Enoch, 2009:321–323.

Aunque también debe mencionarse la versión del constructivismo kantiano de Onora O’Neill. Su reconstrucción del constructivismo parte de la idea de que los agentes morales son finitos, pero también racionales. El razonamiento práctico tendría por objeto llegar a principios que coordinen la acción y la comunicación de esos agentes racionales necesariamente interdependientes. Como cabe esperar en la mayoría de las versiones kantianas del constructivismo, la razón debe ser universal para poder determinar esos principios. Cfr. O’Neill, 1989.

dad de preguntarnos por las razones a partir de las que actuamos o formamos creencias. En el campo práctico en particular, este tipo de reflexión hace posible que los agentes racionales desarrollen ideales acerca del tipo de personas que quieren ser y, sobre todo, de autogobernarse a partir de estándares universales. Ahora bien, la manera adecuada de autogobernarse no es otra que la autolegislación kantiana, que se basa en principios universales. Solo la autolegislación por medio de principios universales representa un verdadero acto de autonomía antes que mero móvil basado en preferencias o deseos irreflexivos y fundamenta la integridad que Korsgaard atribuye a los agentes racionales.

Además, si los agentes morales valoramos algo, prosigue el argumento, esto debe estar fundado en alguna razón moral o valorativa. Pero la idea de que toda valoración por parte de los agentes morales requiere una razón que le dé sustento lleva a una eventual regresión al infinito. Obviamente la regresión al infinito resulta una alternativa inaceptable. Para evitarla, es necesario contar con un valor que pueda dar servicio de razón para fundar el resto de nuestras actitudes valorativas y que a su vez no requiera ninguna razón que la funde a ella misma. Ese valor sería, finalmente, la fuente de la normatividad. Fiel al espíritu kantiano que anima su argumento, Korsgaard afirma que este valor que es fuente de la normatividad es la humanidad. Más aún, es el único valor que puede ser fuente de la normatividad. Así, de la actitud básica de asignar valor que exhiben los agentes morales, Korsgaard encuentra como condición necesaria de esa actitud al valor de la propia humanidad, que obra como contenido formal del principio de universalidad.

Finalmente, la filósofa sostiene que los agentes morales asumen «identidades prácticas» que especifican distintos roles como fuentes de obligaciones especiales.<sup>8</sup> La adopción de una u otra identidad práctica no es, sin embargo, una cuestión caprichosa o carente de fundamento. Más bien, los distintos roles solo asumen la función de identidades prácticas en la medida en la que los agentes adhieren a ellos de manera reflexiva y acorde al principio universal que valora la humanidad del propio agente y de todos los demás agentes. Solo así se asegura la integridad propia de los agentes racionales.<sup>9</sup>

Una variante teórica, distinta a la kantiana, que tiene numerosos defensores es la del constructivismo humeano. Comparte con la versión kantiana del

---

8 Antes de este paso, el argumento de Korsgaard podría ser objeto de una crítica usual a las teorías morales universalistas. Se trata de la objeción según la cual este tipo de teorías no son capaces de explicar de manera acabada la intuición de que los agentes morales tienen obligaciones especiales. El concepto de «identidad práctica» bloquea esa crítica y, en conjunto con la idea del valor universal de la humanidad, evita también la posibilidad de concebir un agente moral que asuma una identidad moral perversa pero perfectamente coherente como la del argumento del mafioso de Gerald Cohen. Cfr. Cohen, 1996.

9 Cfr. Korsgaard, 1996:7–167. Para una reconstrucción crítica puede consultarse Street, 2012.

constructivismo, por supuesto, un compromiso mínimo en la idea de que «la verdad moral no excede el que esta se derive del punto de vista práctico» (Street, 2010:371). Sin embargo, se distingue de ella porque rechaza la idea de que exista algún componente en la razón, ya sea su universalidad o el valor de la humanidad, que sea necesariamente común al punto de vista práctico de todos los agentes morales. En referencia al argumento antes descrito de Korsgaard, Sharon Street rechaza la idea de que la actitud de atribuir valor que despliegan los agentes morales requiera a su vez un compromiso con la idea de que los agentes morales, o la humanidad, son por esa misma razón valiosas. Si se trata de dar razones para las atribuciones de valor de los agentes morales, de acuerdo con los constructivistas humeanos, lo mejor a lo que podemos aspirar es a llegar a una «red coherente de valores interconectados» que fundamenten, de manera contingente, nuestras actitudes valorativas particulares (Street, 2012:51). De modo que el principal contraste entre los constructivismos kantiano y humeano resulta en el hecho de que mientras el primero afirma que del punto de vista práctico de la razón se siguen ciertos compromisos morales objetivos y universales, el segundo lo niega. A modo de ejemplo, quien defienda el constructivismo kantiano rechazaría que pudiera calificarse como un agente racional a un personaje como Calígula o a quien asumiera la «identidad práctica» de un mafioso, dado que obtienen placer en dañar a terceros y esto atenta contra el valor de la humanidad. Para el constructivista humeano, sin embargo, ambos personajes son perfectamente posibles en la medida en la que representen los puntos de vista prácticos y contingentes de ciertos individuos.

También existen tantas otras variantes menos exploradas de posiciones constructivistas metaéticas globales, entre las que quizás vale mencionar como distintivas a la del constructivismo aristotélico. Ella también afirma que la verdad o validez de los juicios morales depende de algunas características del razonamiento o la agencia práctica. Sin embargo, la diferencia relevante con el constructivismo kantiano reside en que el constructivismo aristotélico rechaza la idea de que esas características relevantes sean procedimentales o formales. Más bien, los principios del razonamiento práctico encuentran su fundamento en una concepción sustantiva de la buena vida, «en la que el buen vivir es el objeto de la construcción» (LeBar, 2008:192).

Como fuera adelantado previamente, el constructivismo moral también se caracteriza por un conjunto de posiciones que defiende el constructivismo no tanto como una tesis metaética acerca de qué hace verdaderos o válidos los juicios morales, sino como una posición normativa acerca de la fuente de validez de los juicios morales que debemos aceptar. En este sentido, evitan pronunciarse acerca de tesis globales respecto de aquello que hace verdaderos o



falsos a los juicios morales. La tesis normativa en cuestión puede enunciarse como aquella que afirma que debemos aceptar los juicios o principios morales que se deriven de un procedimiento adecuado de construcción. La mayoría de estas posiciones, a su vez, restringen la validez de su «constructivismo normativo» a ciertas esferas particulares del dominio de lo práctico.<sup>10</sup> Quizás los ejemplos más famosos de estas posiciones constructivistas normativas y restringidas sean el constructivismo acerca de los juicios acerca de lo correcto y lo incorrecto de Thomas Scanlon y el constructivismo acerca de los principios de justicia política y social de John Rawls.

En el caso de Scanlon, su constructivismo se aplica a la moralidad individual de «lo que nos debemos los unos a los otros» (2000; 2012). Este filósofo norteamericano afirma que la corrección de una acción depende de que no se encuentre prohibida por un principio que nadie podría rechazar razonablemente. En su teoría, el «rechazo razonable» es el procedimiento relevante para determinar el tipo de juicios morales que podemos aceptar en la moralidad individual. Por su parte, Rawls afirma que los principios de justicia social y política que debemos aceptar son aquellos que serían elegidos en la situación hipotética de deliberación conocida como la «posición original». Aunque el autor de *A Theory of Justice* siempre dejó en claro que el argumento de la «posición original» se encontraba restringido al desarrollo de principios de justicia social y política, la concepción de su particular dispositivo constructivista no fue siempre la misma. En el antes mencionado artículo *Kantian Constructivism in Moral Theory* de 1980, claramente asociaba el argumento de la «posición original» a un constructivismo de tipo kantiano basado en la autonomía. Sin embargo, en la última exposición de su posición, en *Political Liberalism*, Rawls la reconstruye de manera distinta. En rigor, caracteriza su constructivismo como uno «político» que se articula como un ejercicio de la razón práctica a partir de una «una compleja concepción de la persona y la sociedad»<sup>11</sup> y «especifica una idea de lo razonable y la aplica a sus distintos sujetos» (Rawls, 2005:93-94).

Finalmente, podría atribuirse a David Copp una forma de constructivismo metaético restringido. Su posición parece defender un «constructivismo

---

10 Para caracterizar a las variantes restringidas y normativas del constructivismo resulta elocuente la afirmación de Street según la cual «las perspectivas constructivistas restringidas dan una explicación de la verdad de un conjunto de juicios normativos apelando a lo que se deriva del punto de vista de quien ya acepta (o cuya situación supone la aceptación de) la verdad de algún otro conjunto de juicios normativos. Por esta razón, las perspectivas constructivistas *restringidas* son en última instancia compatibles con cualquier posición metaética» (Street, 2010:368).

11 Debemos mencionar que la inclusión de ciertas premisas fácticas respecto a la sociedad en el dispositivo constructivista rawlsiano fue objeto de una potente crítica por parte de G.A. Cohen. Cfr. Cohen, 2008:229-343.

social» y afirma que los estándares morales verdaderos son el resultado de un procedimiento de decisión que incorpora las necesidades y valores de la sociedad y sus circunstancias. Así, la verdad moral se construye de acuerdo con lo que sería racional elegir para las sociedades. Quizás lo que distingue a esta posición de los constructivismos kantianos es que no apela en ningún sentido sustantivo al valor de la autonomía como garante de lo autoritativo del razonamiento práctico.<sup>12</sup>

### **El problema de Eutifrón y la estrategia constitutivista**

La objeción más frecuente que suele enfrentar el constructivismo es sus diversas formas es una suerte de paradoja similar a la que describe Platón en *Eutifrón*. En dicho texto Sócrates le pregunta a su interlocutor si «¿Acaso lo pío es querido por los dioses porque es pío, o es pío porque es querido por los dioses?» (Platón, 1985:231, 10a). Podría decirse que en ese fragmento el dilema concierne a si los dioses quieren lo valioso porque es valioso independientemente de ellos o bien si lo valioso es tal por ser querido por los dioses. Si se diera la primera opción, los dioses no serían omnipotentes porque hay algo que confiere valor distinto a ellos. En cambio, dada la segunda opción, lo valioso dependería de la voluntad arbitraria de los dioses. Trasladado ese problema al objeto de este capítulo, la objeción plantea un dilema al constructivismo de la siguiente manera. O bien el procedimiento constructivista está fundado en algún otro valor independiente, o bien no lo está.<sup>13</sup> Esta última opción parece indeseable en la medida que implica que el constructivismo es arbitrario. Pero la primera opción tampoco es prometedora para el constructivismo porque parece comprometerlos con la idea de que hay algo independientemente valioso que fundamenta al procedimiento constructivista. Evaluemos, sin embargo, las distintas respuestas que podría dar un defensor del constructivismo según se comprometa con su variante normativa o metaética.

En el primer caso, quien defienda una variante normativa y restringida a una esfera del dominio de lo práctico podría morder la bala de la objeción y tomar el primer cuerno del dilema. Después de todo, este tipo de constructivismo solo afirma una tesis normativa que tiene un alcance limitado. Esa tesis podría estar desde luego fundada en alguna otra «consideración

---

12 Curiosamente, David Copp, el principal defensor de la variante social del constructivismo, también afirma que esta posición es compatible tanto con el realismo como con el naturalismo. Cfr. Copp, 1995. Con todo, semejante compatibilidad sugiere que, en el plano metaético, el constructivismo podría ser eliminable.

13 Cfr. Bagnoli, 2013a:14.

normativa» independiente (Enoch, 2009:333). Por ejemplo, podría sostenerse que el procedimiento constructivista respecto a los valores que debemos aceptar en la política o en la moralidad interpersonal está fundado en el valor independiente de la imparcialidad. Más aún, como hemos señalado más arriba, esta posición es de hecho compatible con cualquier posición metaética. Algo más complicada es la situación de quien sostenga un constructivismo normativo de alcance global. Aquí obviamente sería imposible recurrir a consideraciones normativas ajenas al procedimiento constructivista para fundamentarlo normativamente. Esto suscita la apariencia de que o bien el constructivismo normativo global es efectivamente arbitrario y no puede explicar qué lo fundamenta, o bien requiere una fundamentación moral independiente que en realidad demuestra su alcance restringido. En definitiva, según esta última opción hay valores o principios morales que no son construidos por el mismo procedimiento y por ende este tipo de constructivismo no es global como lo pretende.

La variante metaética del constructivismo es la que enfrenta mayores problemas frente al dilema. Más aún, el problema afecta tanto a su variante global como a la variante restringida. En efecto, aquí vuelve a postularse el dilema acerca de qué es lo que fundamenta la tesis constructivista metaética. Si no ofrece ningún fundamento alguno, resulta arbitraria. Esta alternativa es inaceptable. Sin embargo, si se ofrece algún fundamento en su defensa, o bien este fundamento es el resultado de algún otro procedimiento de construcción o bien no lo es. La primera alternativa solo reproduce el dilema antes planteado, con lo cual es necesario considerar la segunda. Esta, que el constructivismo metaético tenga fundamento en algo no construido, parece comprometer al constructivismo con alguna forma de realismo y los problemas asociados a esta posición.<sup>14</sup> En particular, esta es una objeción que se formula a la mayoría de los defensores del constructivismo que asocian el «punto de vista práctico» con algún elemento de alcance universal como la autonomía o la racionalidad para garantizar la objetividad del constructivismo. Aunque esto se atribuye originalmente a la apelación a la conciencia de la ley moral como un «hecho de la razón» de Kant, también es recurrente en ciertos constructivismos contemporáneos, como el de Korsgaard (Kant, 2011:36, v:32; Korsgaard, 1996:90-130; Street, 2012:48-50).<sup>15</sup>

---

14 Es decir, al menos el problema de la «rareza» ontológica de los hechos morales si no se reducen a hechos naturales, y el problema motivacional derivado de pretender actuar a partir del conocimiento de ciertos hechos morales.

15 Como se señaló previamente, la posición constructivista humeana evita este tipo de suspicacias realistas negando el hecho de que del punto de vista práctico se sigan obligaciones morales universales. En buena medida se podría afirmar que renuncian a la pretensión de objetividad del constructivismo para evitar ese problema.

Para evitar el aparente compromiso realista de ese cuerno del dilema, la mayoría de las defensas contemporáneas del constructivismo adoptan una estrategia constitutivista. Ella consiste en afirmar que el constructivismo no está fundado en consideraciones realistas, pero tampoco es arbitrario, sino más bien que es constitutivo del «punto de vista práctico» involucrado en la agencia moral.<sup>16</sup> De acuerdo con esta manera de entender el constructivismo, el «punto de vista práctico» no requiere ni depende de fuentes externas. Al contrario, construye sus propios estándares de corrección internos que lo hacen una fuente «auto-validante» de valores y normas morales.<sup>17</sup> La objetividad, de acuerdo con esta forma de entender la posición, se deriva de los propios criterios internos y constitutivos de la racionalidad práctica que describe el constructivismo.

Con todo, aun esta manera de defender el constructivismo enfrenta importantes objeciones. Estas trascienden el debate entre kantianos y humeanos respecto a si ciertos elementos, quizás los formales, del punto de vista práctico son universales, si se requiere un valor más sustantivo para fundamentar ello, o bien si no tienen alcance universal y son contingentes. En particular, por un lado, se critica al constitutivismo que es «incapaz de explicar la acción incorrecta» (Enoch, 2009:330). En la medida en que los estándares morales son constitutivos del «punto de vista práctico», parece imposible que un agente pueda violarlos. Si eso fuera posible, se podría señalar que ese agente no comparte el «punto de vista práctico» y por ende, más bien, no es un agente moral.<sup>18</sup> También se ha sugerido que, al contrario de lo que afirman los constitutivistas, la agencia práctica no es una actividad inescapable sino opcional y que por ende no puede fundar la normatividad sin razones independientes para ello (Enoch, 2006; 2009). Finalmente, Michael Bratman también ha puesto en duda la capacidad de la estrategia del constructivismo constitutivista para resolver lo que él denomina «el problema de la alineación» (Bratman, 2012:81). Se trata del problema de la consistencia entre las razones normativas que los agentes de hecho tienen y los requisitos formales postulados como constitutivos de la agencia práctica. En la medida en la que los segundos impongan restricciones muy exigentes a las razones y compromisos normativos que los agentes de hecho tienen, más difícil resulta que el constructivismo explique nuestra vida moral ordinaria.

---

16 Solo para mencionar algunas defensas del constructivismo que apelan a la estrategia constitutivista cfr. Korsgaard, 2009:5; Bagnoli, 2013a:2; 2013b:154; Street, 2008:55; 2010:373.

17 Cfr. O'Neill, 1989; Korsgaard, 2008; Street, 2010.

18 Para una crítica similar véase Cohen, 1996:174–177.

## Conclusión

Como se ha intentado demostrar a lo largo de este capítulo, la del constructivismo es una posición articulada y potente que se ha ganado un lugar en la filosofía moral actual. En buena medida, su principal atractivo consiste en pretender detentar las respectivas ventajas del realismo y el antirealismo sin sus problemas.

Asimismo, el constructivismo demuestra un enorme potencial de desarrollo prospectivo. Todavía no es posible encontrar un consenso unánime, entre otras cosas, pero principalmente, en torno al modo correcto de concebir y fundamentar la fuerte objetividad que promete esta posición sin el recurso a «hechos morales» del realismo. En todo caso, la respuesta humeana a este problema parece más bien, como se mostró más arriba, una retirada antes que una solución al mismo.

Con todo, puede incluso pensarse que lo más interesante del constructivismo no se encuentra en la respuesta que pueda ofrecer a los problemas tradicionales de la metaética y la filosofía moral. Más bien, puede ser el punto de partida para desarrollar cabalmente la intuición que se advierte detrás de él. Se trata de la sugerencia de deponer la aproximación teórica o cognitiva aplicada en filosofía moral y reemplazarla en cambio por una orientación marcadamente práctica. Los alcances y límites explicativos de la vida moral por parte de esta forma de entender la filosofía práctica sin duda merecen ser evaluados, así como su supuesta capacidad distintiva para guiar la acción.

## Referencias bibliográficas

- Bagnoli, C. (2013a). Introduction. En Bagnoli, C. (Ed.), *Constructivism in Ethics* (pp. 1–21). Cambridge University Press.
- Bagnoli C. (2013b). Constructivism about practical knowledge. En Bagnoli, C. (Ed.), *Constructivism in Ethics* (pp. 153–182). Cambridge University Press.
- Bratman, M. (2012). Constructivism, Agency, and the Problem of Alignment. En Lenman, J. y Shemmer, Y. (Eds.), *Constructivism in Practical Philosophy* (pp. 81v98). Oxford University Press.
- Cohen, G.A. (1996). Reason, humanity and the moral law. En Korsgaard, C., *The Sources of Normativity* (pp. 167–188). Cambridge University Press.
- Cohen, G.A. (2008). *Rescuing justice and equality*. Harvard University Press.
- Copp, D. (1995). *Morality, Normativity and Society*. Oxford University Press.
- Darwall, S.; Gibbard A. y Railton, P. (1992). Toward Fin de siècle Ethics: Some Trends. *The Philosophical Review*, 101(1), 115–189.
- Dworkin, R.M. (1973). Justice and Rights. En Dworkin, R.M., *Taking Rights Seriously* (pp. 150–183). Harvard University Press.
- Enoch, D. (2006). Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Agency. *Philosophical Review*, 116(2), 169–198.
- Enoch, D. (2009). Can there be a global, interesting, coherent constructivism about practical reason? *Philosophical Explorations*, 3(12), 319–339.
- Kant, I. (2012). *Fundamentación de la Metafísica de las Costumbres*. Las Cuarenta.
- Kant, I. (2011). *Crítica de la razón práctica*. Fondo de Cultura Económica.
- Korsgaard, C. (2008). Realism and Constructivism in Twentieth-Century Moral Philosophy. En Korsgaard, C., *The Constitution of Agency* (pp. 302–326). Oxford University Press.
- Korsgaard, C. (1996). *The Sources of Normativity*. Cambridge University Press.
- LeBar, M. (2008). Aristotelian constructivism. *Social Philosophy and Policy*, 25(1), 182–213.
- Mackie, J.L. (1988). *Ethics: Inventing Right and Wrong*. Penguin.
- Nino, C.S. (1988). *Constructivismo epistemológico: entre Rawls y Habermas*. Doxa.
- Nino, C.S. (2007). *Ética y derechos humanos*. Astrea.
- O'Neill, O. (1989). *Constructions of reason*. Cambridge University Press.
- Platón (1985). Eutifrón. En *Diálogos I*. Gredos.
- Rawls, J. (1980). Kantian Constructivism in Moral Theory. *The Journal of Philosophy*, 77(9), 515–572.
- Rawls, J. (2005). *Political Liberalism*. Columbia University Press.
- Scanlon, T. (2012). The Appeal and Limits of Constructivism. En Lenman, J. y Shemmer, Y. (Eds.), *Constructivism in Practical Philosophy* (pp. 226–242). Oxford University Press.
- Scanlon, T. (2000). *What We Owe To Each Other*. Harvard University Press.
- Street, S. (2010). What is Constructivism in Ethics and Metaethics? *Philosophy Compass*, 5(5), 363–384.
- Street, S. (2012). Coming to Terms with Contingency: Humean Constructivism about Practical Reason. En Lenman, J. y Shemmer, Y. (Eds.), *Constructivism in Practical Philosophy* (pp. 40–59). Oxford University Press

### 3. Realismo moral

*Guillermo Lariguet*

#### **Para empezar: algunas citas filosóficas para calentar los «músculos filosóficos»**

Everything is what it is and not another thing. (Butler, 2006:39)

O como Plotino escribe: «Lo que es no difiere de lo que es, igual que un filósofo no difiere de un verdadero filósofo (...) Nosotros añadimos a «lo que es» la palabra «siempre», y a «siempre» la palabra «ser», y nos referimos entonces al «ser que dura siempre». Lo cual significa: «lo que es siempre, es lo que es verdaderamente». (Arendt, 2001:30)

Quien tenga una creencia acerca del mundo —o acerca de cualquier otra cosa— debe captar el concepto de verdad objetiva, de lo que es el caso independientemente de lo que él o ella piensa. Debemos preguntar, por tanto, por la fuente del concepto de verdad objetiva. (Davidson, 2003:285)

Los poderes deónticos tienen una característica única que, de nuevo, considero extraña y quizá desconocida en el mundo animal: una vez reconocidos, nos proporcionan razones para actuar, que son independientes de nuestras inclinaciones y deseos. (Searle, 2017:25)

La tesis que voy a sostener es la siguiente: la crueldad del coma punitivo no depende de lo que consideraban los constituyentes, ni de lo que considera ahora la sociedad, ni los jueces del Tribunal Constitucional. Como dijo H. Putnam «los significados no están en la cabeza». Y, por lo tanto, no es el significado lo que determina la referencia de nuestras expresiones, sino precisamente al revés: es la referencia lo que determina el significado. (Moreso, 2010:5-48)

«Todo es lo que es, y no otra cosa», dice, en vena tautológica, el Obispo y filósofo del siglo XVIII Joseph Butler, a quien G.E. Moore (1997) cita en la célebre *Principia ethica*; cualquier creencia sobre el mundo presupone la verdad objetiva, dice Davidson, ya en el siglo XX, «de lo que es el caso, independientemente de lo que pensamos». Lo que es el caso independientemente de cómo lo pensemos, escribe Plotino, «es idéntico a sí mismo siempre» y captarlo está entre las condiciones que nos permiten aspirar a ser verdaderos

filósofos. Los poderes deónticos (las normas que nos obligan, prohíben o permiten conductas), afirma Searle, nos dan razones para actuar que «son independientes de nuestras inclinaciones y deseos». Categorías morales, como la crueldad del coma punitivo, idea de sanción extraída por el autor de estas líneas, el filósofo contemporáneo José Juan Moreso, de la novela de Philip Kerr *Una investigación filosófica*, «no dependen de lo que consideraron los constituyentes» (los padres de una Constitución). La ontología de la crueldad tampoco depende, dice Moreso, de lo que opine la gente, lo que, habitualmente, llamamos moral positiva. Y esta opinión se extiende incluso a los jueces. Es la «referencia, afirma el filósofo catalán, no al revés, lo que determina el significado», en este caso, el significado de categorías morales.

¿Qué tienen que ver estas afirmaciones, aparentemente diversas en naturaleza filosófica, entre sí? ¿Es un disparate conectarlas? Por sobre las eventuales obsesiones por evitar puentes salvajemente anacrónicos entre tesis diversas en sus localizaciones temporales, una pregunta genuinamente filosófica sería esta: ¿qué enlaza a un filósofo del siglo III después de Cristo, pasando por uno moderno como Butler, con afirmaciones de filósofos contemporáneos como Davidson, Searle o Moreso? Mi respuesta es que hay una «comunalidad parcial» subyacente a estas afirmaciones; lo que podría llamarse su «espíritu subyacente» descansa en una doctrina filosófica que viene de muy lejos en el tiempo: me refiero al realismo en general, y al realismo moral en particular. La sustancialidad de este espíritu encarna en una intuición filosófica como la siguiente: hay aspectos del mundo en general, del mundo moral en particular, que no dependen, en un sentido importante que me propongo dilucidar en este capítulo, de lo que pensemos, o deseemos.

Desde luego que, dicho lo anterior, me veo impelido a efectuar una aclaración previa importante. Pues, la postulada independencia del mundo, respecto de creencias o deseos, presupuesta en el realismo *toto genere*, tiene matices distintos según el tipo de postura que esté en juego en cada filósofo.

Así, la comunalidad de la que acabo de hablar, para definir la sustancia del realismo moral, no es total, sino parcial como sostuve más arriba. Plotino es un metafísico preocupado por identificar una realidad última inderivable, *lo Uno*; Butler es teólogo y sus propósitos también son transmundanos. Davidson cree en la verdad objetiva en un sentido específico que, en filosofía de la mente, llamamos verdad de la «tercera persona». A su vez, para Davidson, dicha verdad debe ser lógicamente consistente con la verdad de la llamada «primera persona», que se conecta con la autoridad que tenemos, en principio, para dar cuenta de nuestras (propias) creencias o percepciones subjetivas. Y, finalmente, el mundo de la tercera persona se conecta con el de la primera, por el expediente de lo que Wittgenstein vio en sus *Investigaciones filosóficas*:



la verdad de la «segunda persona» enraíza en nuestras relaciones intersubjetivas, de los que, como nosotros, según el filósofo vienés, compartimos un mundo de la vida, una práctica, un lenguaje, unas reglas, una esfera pública. Así, el yo de la primera persona, al carecer de un lenguaje estrictamente privado, para Davidson, y para Wittgenstein a quien el primer nombrado sigue en esto, no es solipsista y se ensambla con una visión objetiva del mundo (imparcial, desde un punto de vista tercerizado), a través del «tú», del «otro», de la intersubjetividad. La realidad así, para Davidson, está triangulada, y los tres mundos son, en rigor lógico, un solo mundo. La objetividad, sin embargo, no involucra un substrato metafísico «trascendente», sino uno inmanente al mundo de la vida y al lenguaje compartido.

Algo similar a lo anterior, parece seguirse del proyecto filosófico de Searle. Para él no hay dos mundos (el físico y mental), ni tres, como decía Sir John Eccles (el físico, el mental y el cultural) sino un solo mundo. Para Searle, constituye un requerimiento básico de autocomprensión analítica, que nuestra reconstrucción del mundo normativo, o del moral para lo que será mi elucidación presente, dependa de hechos muy básicos sobre el mundo: hechos que son físicos (la realidad del átomo), químicos (la realidad del enlace químico) y biológicos (la realidad del ADN).

La realidad de la que quiere dar cuenta Serle, que es de naturaleza «social», específicamente normativa, en un sentido institucional singular, tiene que poder explicarse, sin cometer reducción lógica, en las propiedades más básicas del universo. Por lo que, para Searle, más allá de las distinciones, por ejemplo, entre tres mundos de Eccles, o de niveles (subjetivo, intersubjetivo, objetivo) de Davidson, hay un solo mundo. Esta afirmación, como se verá en este capítulo, no carece de consecuencias filosóficas para la discusión del alcance y posibilidades lógicas del realismo moral. Para John Searle, las «razones para actuar», son *independientes* de deseos o inclinaciones de los agentes humanos. Y en virtud de tal independencia postulada, sí que las mentadas razones asignan *objetividad* al razonamiento práctico. Ahora bien, se trata de una objetividad con un trasfondo ontológico peculiar: la constatada a partir de «instituciones». Para Searle las mismas dependen de una mentalidad (intencionalidad) colectiva convergente, que se trasunta en un lenguaje declarativo uniforme, que representa, a su tiempo, funciones de estatus social (por ejemplo, ser padre, marido, juez, propietario, jugador, asociado, votante, etc.). Dichas funciones se articulan con poderes deónticos (o sea, lo que podemos hacer conforme reglas). Y la *realidad normativa*, en la que medita Searle, que es la que me interesa por su proximidad con el tema del capítulo, el realismo moral, es *creada* por reglas —constitutivas— que, en uso de un lenguaje conceptual, articulamos. Sin esta creación —subjetiva ontológica—

mente— no habría una realidad epistémicamente objetiva: la institucional. Esta realidad «creada» tampoco es la realidad moral, *stricto sensu*, de la que habla el realista. O si lo fuera, requeriría de afirmaciones, matices y detalles que, quien quisiera argüirla, *concesso non dato*, debiera hacerlo tras una larga clarificación.

Por último, Moreso alude a la teoría causal de la referencia de Kripke–Putnam (contrapuesta a la teoría descriptiva de Frege–Russell) conforme la cual el significado, como proposición mental es independiente de nuestros usos lingüísticos, de nuestras convenciones, (o realidad institucional como dice Searle). Es la referencia a la «causa» del significado, no al revés lo que debe primar en el análisis filosófico. Esta forma de pensar está ya más cerca del hábito del realismo moral.

No obstante las digresiones que se acaban de hacer, se puede aseverar, a pie juntillas, que las citas del epígrafe de este capítulo, sí que captan elementos que, normalmente, podríamos estar dispuestos a adscribir *parcialmente* al realismo moral como categoría del pensamiento metaético. Prueba de lo dicho es lo que ya se ha indicado con las citas filosóficas del comienzo: las cosas son como son, las creencias pueden ser verdaderas o falsas porque hay hacedores de verdad ahí afuera que objetivamente lo permiten, los poderes deónticos dan razones para actuar independientes de nuestras inclinaciones habituales o contingentes y el significado moral está determinado por el mundo, por las referencias pertinentes, no por nuestras descripciones. Pero digo que las conexiones son parciales por cuanto no necesariamente algunos de los citados, como Davidson, por ejemplo, no por tematizar la cuestión de la objetividad se vuelve, *ipso facto*, un realista moral. Después de todo, ser objetivista y ser realista son posturas que pueden y deben distinguirse. Un realista es un objetivista, pero no todo objetivista necesita ser un realista y menos uno de índole moral.

## **El realismo moral: primeras precisiones conceptuales**

La variedad conceptual reseñada líneas atrás, junto a la señalada comunalidad parcial de las citas filosóficas de la sección anterior, sus matices y direcciones conceptuales proyectadas, sugieren que el tema del realismo moral trasunta una notoria complejidad filosófica.<sup>1</sup>

---

1 Precisamente, una de estas complejidades, que aquí he decidido no tratar por motivos de auto-restricción metodológica, está vinculada a lo que se conoce actualmente con el nombre de «nuevo» realismo, o realismo «especulativo», u otras denominaciones. Se trata de una literatura no específicamente vinculada al realismo «moral» y que tiene raíz más bien «continental», que analítica (véase, sólo a título de ejemplo, Ferraris, 2012). En trabajos independientes a éste sería relevante

Para comenzar, sostendré que el realismo moral es una especie dentro de un conjunto más vasto de asunciones generales sobre el universo. En segundo lugar, el realismo moral es una doctrina de carácter metaético.<sup>2</sup> Ello significa que, en principio, se asume que el realismo moral es un conjunto de tesis que hacen parte de la ética que denominamos «filosófica».

La ética filosófica, o metaética, es una creación de la filosofía analítica del siglo xx. Se supone que la metaética es filosófica en un sentido muy estricto: sus tesis, cualesquiera sean, se dirigen a pensar los presupuestos (ontológicos, metafísicos, semánticos, epistémicos) de los problemas filosófico-morales y no a emitir pautas normativas acerca de cómo actuar o cómo forjarse un carácter moral. Esto hace que la metaética difiera de la ética *sustantiva* o «normativa». Esta última procura dilucidar conceptualmente los requisitos que hacen posible o imposible la acción moral. Por ejemplo, es propio de la ética sustantiva o normativa reflexionar sobre el modo en que las máximas de la acción (las inclinaciones subjetivas) pueden transformarse, de modo consistente con todo el universo normativo de la moral, en normas morales válidas u objetivas. Esta es la motivación que tiene Kant para urdir el imperativo categórico. La ética normativa puede asumir otros rostros: puede abrazar alguna forma de ética de la virtud si, como por ejemplo con Aristóteles, se busca definir qué hábitos hacen de un carácter personal, un buen carácter moral. O, también, puede asumir una versión consecuencialista como la que Bentham propagó con su obra sobre el utilitarismo. Aquí, se procura establecer, normativamente, una suerte de métrica que permita identificar qué cursos de acción reporta mejores consecuencias o, en términos de un principio llamado *maximin*, evitar las peores consecuencias.

Se ha discutido, y se discute todavía, si la metaética es conceptualmente posible. Algunos filósofos se sienten genuinamente tentados a defender que no, que toda metaética colapsa, después de todo, en una ética sustantiva o normativa. Aquí apartaré este problema y mantendré, a efectos de, por lo menos, una mera nomenclatura terminológica, que el realismo moral es una doctrina *metaética*.

La tercera precisión que quiero hacer es que el realismo moral, en principio, se distingue del «realismo científico». El realismo científico defiende la doctrina conforme la cual el mundo a conocer (*explicatum*) es *independiente* de la mente explicativa (*explanans*). Puede, por tanto, ser descripto y explicado de manera verdadera o falsa conforme unos hechos crudos o no normati-

---

poner a dialogar esta perspectiva relativamente novedosa con la discusión analítica proverbial en materia de realismo moral.

2 Decirlo así, no impediría defender, a la postre, que hay mayor trabazón lógica, que la que se acostumbra a admitir, entre la metaética, o ética filosófica, o ética de segundo orden, y una ética normativa o de primer orden.

vos. Ahora bien, distinguir el realismo moral encajado en unos hechos más bien normativos, respecto del realismo científico, basado en unos hechos más bien crudos, no es lo mismo que desconectarlos. Como discutiré oportunamente, hay formas científicas que puede asumir el realismo moral; formas que, adelantando un paso aquí, diré que se denominan «naturalistas». Y, dicho sea de paso, aunque el realismo moral no se considere científico, o naturalista, no implica que abandone algunos postulados estrella del propio realismo en general, y del científico en particular, a saber:

- a) Que hay un conjunto de proposiciones básicas —sobre el universo físico, por ejemplo— que pueden ser, y que son, verdaderas.
- b) Que las propiedades descritas o explicadas por aquellas proposiciones básicas (que forman su contenido) son independientes de la mente explicativa o cognoscente.

La cuarta y última distinción que plantearé en este apartado se vincula con la idea según la cual el realismo moral eleva *tesis* de diferente factura. Solo a título de ejemplo, quiero mencionar estas:

- a) Existe una realidad que puede ser conocida por los agentes humanos por diversos métodos: introspección, observación, percepción, intuición, reflexión, etcétera.
- b) Dicha realidad es, en algún sentido, incluso si tiene rasgos naturalistas, sui generis. Ello porque, entre otras cosas primordiales, es una realidad *normativa* de algún tipo; una realidad «moral» que debe —y puede— discernirse de una realidad científica.
- c) Conforme b), el objeto nuclear del realismo moral es la realidad moral. Tal realidad, por definición intuitiva de la expresión «realidad», está constituida por «hechos», pero no los llamados hechos «brutos» (o básicos) de los que se ocuparía, por caso, la física. Sino por hechos morales cuya relación con los hechos brutos o básicos hay que explicar también.
- d) Los mentados hechos morales son, en diversos sentidos que deben ser despejados, *independientes* de la «mente moral», la cual se edifica a partir de un conjunto complejo de procesos que involucran creencias, intenciones, deseos y voliciones.
- e) La referida independencia estaría en condiciones de asegurar una noción *objetiva* de verdad. En términos tarskianos, una noción de verdad *qua* correspondencia. Una proposición —moral— o científica, es verdadera si y solo si su contenido se «ajusta» al mundo tal como es.
- f) Las proposiciones pueden entenderse, desde la filosofía del lenguaje, y la lógica proposicional, como el «significado» de términos, en este caso morales.

- g) Tal significado es establecido por la referencia y no al revés. Esto hace que el realismo moral se lleve mejor, en principio, como alega José Juan Moreso en una de las citas de la sección anterior, con la teoría causal de la referencia de Kripke–Putnam, que con la teoría descriptivista de Frege–Russell.
- h) Por último, se supone que un realista moral debe admitir, por mor de su doctrina, que los hechos morales —cualquiera sea su naturaleza conceptual— otorgan razones para actuar. «Razones para actuar» no en un sentido meramente explicativo–causal, sino en un sentido justificatorio o normativo al cual ajustar la conducta moral.

La tesis a) es *epistémica*. La realidad moral es susceptible de ser conocida, con lo cual, a contrapelo de Kant, el nómeno moral sí que puede ser accedido mentalmente. Ello por diversas vías metodológicas, como las ya enunciadas arriba, cada una con su complicada peculiaridad. Las tesis b) y c) son *ontológicas y metafísicas*. Ontológicas porque refieren a que *existen entidades morales* (hechos, propiedades de hechos). Metafísica porque tales hechos tienen una naturaleza propia que puede ser, por ejemplo, de índole natural o trans–natural, según el caso. En caso de que tal índole sea trans–natural, como opina George Moore, tenemos un problema que se formula preguntándonos cómo es posible conciliar la realidad *sui generis*, que es la realidad moral, con la realidad —no *sui generis*— de la realidad científica; o sea, cómo es posible hablar, de manera lógicamente defendible, de lo que Searle quiere conjurar con su teoría: la existencia de dos mundos, en vez de uno. El otro problema es que la mente moral parece asumirse como un conjunto de procesos distanciados de los hechos morales. Esta «distancia» es el requisito, a veces tácito, que permite que los hechos puedan ser asumidos como independientes. Sin embargo, es un dato constatable, y por demás razonable, que la mente moral también es una forma de hecho, y de hecho moral si se quiere. Con lo cual, si no se tiene en claro, primero, la naturaleza de la mente en general, de la mente moral en particular, poco se podrá avanzar en postular, ingenuamente, que hay entidades independientes a ser conocidas.

La tesis d) está a caballo de lo epistémico y lo ontológico: por un lado, «existe» (como un dato óptico) la referida «independencia» de los datos morales (independencia que reafirman, con sus matices diversos, las citas de Davidson, Searle o Moreso); por el otro lado, esta independencia es también susceptible de conocimiento (el rasgo epistémico del que acabo de hablar). Esta susceptibilidad de conocimiento moral puede querer decir, por lo menos, dos cosas distintas. O bien, que la realidad moral es «realmente», valga la redundancia, independiente, o bien que tal independencia se asume como, diría el finado Parfit (2011), como prerequisite lógico para conocer ciertas en-

tidades. Algo parecido a cuando presuponemos que los números, o las relaciones lógicas, realmente existen, presuposición efectuada a los fines estrictos del análisis conceptual.

La tesis e) apela al concepto de verdad, verdad moral si se prefiere, que hay que presuponer como parte nodal del realismo moral, cualquiera sea su versión. Está claro que esa noción es la correspondentista tarskiana de vieja prosapia: la metafísica de Aristóteles, por ejemplo, podría ser un ejemplo primario de esa versión. Ahora bien, aunque es poderosamente intuitiva esta noción de verdad, sabemos que aquí hay un atolladero lógico denunciado por Ronald Dworkin (2014) en su obra *Justicia para erizos*. El problema, como lo transmite el fallecido profesor americano, es que es falso, dirá él, que estemos en «contacto causal» con hechos o propiedades morales de hechos. No es cierto, alega Dworkin, que estemos causalmente vinculados a estos hechos (*morons*, los llama irónicamente él) pues, si así fuera, no debería haber *desacuerdo moral*. Si todas nuestras mentes morales están causalmente conectadas con el mundo moral, entonces, ¿cómo es posible encontrar desacuerdos morales, a veces tan profundos? La relación de causalidad, de existir, garantizaría lógicamente que no hubiera desacuerdo pues, todas nuestras «ideas» morales serían causadas de manera homogénea y uniforme por esta presunta realidad moral. El atolladero, identificado por Dworkin, tiene un aire de familia con un argumento que, normalmente, los éticos como Leandro Paolicchi en este mismo libro, llamamos «escéptico» sin más. Y es el argumento fraguado por John Mackie (2000), en su obra *Ethics. Inventing right and wrong*. Para Mackie, en efecto, un «signo» de la ausencia sistemática de hechos morales es que tenemos reales desacuerdos morales. Por tanto, es un «error» suponer que los hechos morales objetivos existen. Nuestras creencias, al respecto, son falsas. Es sugerente que el subtítulo de la obra de Mackie use el verbo «inventar». Pues, como Searle, que emplea el vocablo «crear», Mackie admite que hay una suerte de realidad moral más bien *convencional*. Así, una implicancia epistémica de lo dicho por Dworkin, Mackie o Searle, es que la realidad normativa en general, la moral en peculiar, no es *descubierta*, como cuando un explorador descubre una nueva montaña, sino *inventada, creada* con cierta finalidad: por ejemplo, maximizar la cooperación social, minimizar el conflicto, etc. Aun así, no se deja en claro en el argumento si tales fines también son inventados o creados, o si ellos forman parte, en algún sentido inteligible de la expresión, de nuestra naturaleza como agentes humanos.

Las tesis f) y g), finalmente, atañen a una dimensión *semántica* del realismo moral. Ello porque se dirigen a elevar a preocupación de primer rango el tema del significado de los términos morales. Como Moreso apunta en el ya citado trabajo, este significado, para el realista, está hecho por las referencias,

no al revés. No se trata de cualquier clase de referentes. Sino que los referentes son adoptados por la psicología humana como «designadores rígidos». La verdad conceptual de que el agua es una sustancia compuesta de  $H_2O$  tiene que ser cierta en todos los mundos posibles que, en principio, satisfagan la descripción de las leyes naturales de nuestro mundo actual. Extrapolando este razonamiento al campo moral, las proposiciones morales más básicas, por ejemplo, la que indica que matar sin legitimación moral alguna, es incorrecto, tendría que ser una proposición eterna, aplicable a cualquier subconjunto social. Ello por definición: pues los referentes operan rígidamente en materia semántica. Desde luego, desde el punto de vista de una moral positiva, esta versión del significado puede ser motivo de perplejidad. En términos de dicha moral, es una verdad —empírica, más que conceptual— que las apreciaciones morales cambian, de modos que van de lo sutil a lo más grueso, en el tiempo y en los subconjuntos sociales. Por lo cual, la noción de designador rígido solo es compatible con una noción genérica de «moral crítica», no de moral positiva. La moral crítica, se presupone analíticamente, debe entenderse como eterna y refractaria al relativismo social y/o cultural tematizado por Lucas Misseri en esta misma obra.

Última, pero no menos importante, h) eleva una tesis sobre el *razonamiento práctico y la motivación moral*. Según dicha tesis, los hechos morales son recipientes de normatividad. Con esta metáfora quiero dirigir la atención hacia lo siguiente: si los hechos morales son normativos ello debe traducirse en que los mismos son la base del razonamiento práctico, específicamente el que denominaríamos razonamiento moral. Que sean la «base», sin embargo, puede implicar una ambigüedad que hay que deshacer ya mismo. Los hechos pueden ser la «base» explicativa de nuestro comportamiento moral o la base motivante justificatoria de cómo *debemos* comportarnos. Ahora bien, la primera idea de «base» pareciera, en principio, tener que ser descartada. Ello porque solo hechos no normativos fungen como categorías explicativas de nuestro comportamiento. Empero, por definición, hechos normativos, como lo serían supuestamente los hechos morales, no pueden meramente explicar el comportamiento. Su pretensión nodal es justificar lo que debemos, o no debemos, hacer, *all things considered* (todas las cosas consideradas).

Entendido el punto anterior, emergen dos inquietantes problemas. El primero es que, con los hechos, incluso si morales, parecemos mantener una relación «causal», no una relación normativa. Esta aserción era parte del argumento de Ronald Dworkin que retraté líneas atrás. Ahora bien, si estamos en contacto causal con unos reputados hechos, parece que la dimensión normativa de los hechos morales deja ser capturada. Y si esto es así, estaríamos violando la *ley de Hume*, conforme con la cual, está vedado pasar de enunciados

«es» (sobre hechos) a enunciados «debe». Aparentemente, este problema sería conjurado por el realista moral, pues este no construye su doctrina a partir de hechos «a secas», o hechos crudos como dice Sol Yuan en este mismo libro, sino hechos morales. Ahora bien, el sintagma «hechos morales» puede tener diversos sentidos, opuestos incluso entre sí. Un sentido podría entenderse a partir de la noción según la cual tales hechos están montados sobre una doble realidad cuyas conexiones lógicas y/o no lógicas hay que esclarecer. La realidad sería doble si se asume, en forma férrea, que la base de la realidad moral *stricto sensu*, son *hechos no morales*, o sea, lo que se viene denominando aquí con el sintagma «hechos básicos». La relación lógica entre esa realidad (física, química, biológica) y la otra, la realidad moral, que es, por compromisos analíticos necesarios, de carácter normativo, podría ser lo que los filósofos denominamos «superveniencia».

La superveniencia sería un nombre posible para algún tipo de relación de *grounding* —como preconiza en este libro Samuele Chilovi— que unos hechos no evaluativos (naturales) respecto de hechos evaluativos (o morales). Por lo pronto, habría otro sentido de hecho moral, conforme al cual, estos hechos, sustancia misma de la realidad moral, son de índole metafísica: pertenecen a una región del valor, a un *eidos normativo*, independiente, en un sentido a explorar, respecto de los hechos no morales o hechos básicos.

El segundo problema con el que nos enrostra la tesis h) es el de la *akrasía* estudiado por Luciana Samamé en el presente libro. Es, por tanto, el clásico atasco para explicar la motivación moral. El realista moral tiene que explicar, todavía, cómo diablos es posible que del hecho moral mismo de poder «reconocer» cognitivamente hechos morales, se siga que el agente cuenta con la motivación adecuada, justa, precisa, para llevar adelante, moralmente, lo reconocido cognitivamente. Como dice Nicolás Alles en el libro, un realista moral tendría que asumir aquí, por caso, una versión *internalista* de la motivación, conteste con la cual las simples «creencias» morales, o sea, los estados de la mente moral puramente epistémicos, mueven a actuar, *eo ipso*, en el sentido correcto trazado por tales hechos. Un escéptico humeano, empleando diversa munición lógica, intentaría cuestionar que las creencias tengan esa aptitud motivacional intrínseca. Por supuesto que, en el marco de ese cuestionamiento, puede empezarse por lo más básico: la creencia moral en cuestión puede ser falsa y conducir a acciones que, vistas todas las cosas, no sean realmente morales. Pero el segundo cuestionamiento es más profundo y viene a sostener lo siguiente: las creencias no mueven a actuar sino, como afirmó Hume en su *Tratado*, que son las pasiones (los deseos) las que tienen esa aptitud motivacional intrínseca constatable. Una réplica a este argumento podría ser que las creencias morales de las que se parte son aquellas que, en el senti-



do categórico kantiano, califican como «racionales». Pero esta réplica, podría contra-atacar válidamente el humeano, implica pedir la cuestión.

En el fondo de esta discusión, creo yo, se encuentran dos modos diversos, para no decir opuestos, de entender la llamada «psicología moral». En el caso kantiano, este es un modo «racional» en un sentido profundo. Los agentes, realmente humanos, son aquellos que vencen sus inclinaciones contingentes, y operan de forma ética y lógicamente consistente. De lograr lo dicho, se convierten en agentes que, desde un punto de vista modal, se tornan *necesarios*. La psicología humeana, al parecer, es distinta y en esto Hume, no habría despertado de este supuesto «sueño dogmático» a Kant. Se trata, en el caso del escocés, de una psicología «empírica» porque se para mientes en los deseos reales que un agente tiene para actuar. Además, la racionalidad estrella en Hume es instrumental (hipotética, en los términos de Kant). La réplica — parcial— a este punto del kantiano, podría ser que, para Hume, no califica como moral cualquier tipo de deseo bruto. Más bien, califican los deseos «serenos»; un término, podría decir el kantiano, que disfraza cierto disciplinamiento que ejerce la razón sobre las pasiones para moldearlas moralmente en función de una emoción central para Hume: la empatía.

A los dos problemas recién tratados, se añadiría un tercero que, de alguna manera, requiere de una lectura adecuada, una que permita ser realmente fiel a Kant.

Como es sabido, Kant es visto, de manera recurrente, como un ejemplo de «antirrealista moral». Ello porque Kant negaría la existencia de hechos morales independientes de la mente moral. Así, Kant, para filósofos como John Rawls o, en habla hispana, Carlos Nino, sería un ejemplo perspicuo de lo que en esta obra Martín Oliveira presenta como un «constructivista moral»: los hechos morales no están ahí dados, fuera del sujeto, sino que son investidos, en un giro copernicano, por el sujeto moral mismo. Cristina Lafont (2002:115-129), sin embargo, ha puesto en duda, a mi entender con argumentos interesantes, que Kant encarne una forma de antirrealismo moral. Ella prefiere, más bien, llamarlo un «realista anómalo» porque, según Lafont, en Kant no importa solo universalizar exitosamente las máximas de acción, sino que tales normas morales, obtenidas por reflexión moral universalizante, sean «realmente», «de hecho», dice ella, morales. Creo que lo que Lafont quiere transmitir es la operatividad de un deslizamiento semántico sutil que va desde la construcción moral a admitir que la misma tiene que ser, de hecho, realmente una construcción moral. Detrás de la búsqueda de consistencia a la que agentes humanos racionales se hallarían orientados, idealmente enfocados hacia un «reino de fines», habría «hechos», muy básicos, que cabría también llamar de «morales». Estrictamente, hechos de la «razón pura práctica».

El imperativo categórico, en este caso, podría interpretarse como emergiendo de hechos de la razón, constatados por agentes racionales humanos en condiciones ideales.

No es el caso, ahora, de defender aquí, en forma articulada, que Kant no es un antirrealista o constructivista, sino un realista anómalo. Más bien, debo avanzar a la siguiente casilla y explicar dos clases diferentes de realismo moral. Veremos, además, si este par de nociones cobijan o no una tercera forma de realidad: la institucional.

### **¿Dos clases específicas de realismo moral? ¿O hay, acaso, un *tertium datur* entre las dos clases específicas de realismo y hecho moral?**

La doctrina del realismo moral puede pensarse como una bifurcación entre dos clases diferentes de realismo moral. En consecuencia, entre dos clases de hechos morales. Por razones estandarizadas en la filosofía que aquí asumiré como propias, denominaré a estas clases como no naturalista y naturalista, respectivamente.

La clase de realismo moral no naturalista podría encontrarse en Platón. La realidad moral, la del bien moral concretamente, moraría en un mundo ideal. Lo que hoy denominamos «hechos morales» serían entes ideales, formas perfectas a la que una realidad moral ordinaria (empírica, o positiva) debería ajustarse todo lo que resulte posible. Los hechos morales serían, para Platón, conceptos abstractos *separados* de la realidad moral ordinaria. Esta separabilidad se verificará, posteriormente, en la Edad Media cuando, en la discusión sobre los universales, se asuma un realismo conceptual según el cual los conceptos universales verdaderos son *ante rem*. Estos entes morales separados de la realidad moral empírica, también, son *universales* (De Libera, 2015). El rasgo de universalidad implica que estos conceptos morales abstractos (de lo bueno, lo correcto, lo bello, etc.) deben aplicarse a griegos y extraños, a presocráticos y socráticos, en fin, a todo el mundo, y sin que el tiempo empírico infecte la eternidad de los conceptos morales. Quiero argüir que esta versión platónica se mantiene, de algún modo, con los matices del caso, en *Principia Ethica* del ya fallecido Profesor de Cambridge, George. E. Moore, miembro descolante del club de los «Apóstoles». Para Moore, a excepción justamente de Platón, y habría que agregar de Sidgwick, buena parte de la filosofía moral (Aristóteles, estoicos, éticos evolucionistas como Spencer, etc.) han pervivido pese a que sus sistemas filosóficos están afectados de una falacia letal: la *falacia naturalista*. Esta falacia consiste en reducir lo moral a lo no moral.

En otras palabras, a reducir lo moral a lo natural. La palabra «natural», sin embargo, merece una breve glosa clarificatoria. Básicamente, para Moore, lo «natural» es definido como aquello que constituye «objeto de las ciencias naturales» (la física, la química, la biología). Un ejemplo clásico es este: reducir, como pretendía, aparentemente Bentham, lo moral a lo que resulta *placentero*. Someter lo moral, lo bueno o correcto, a la soberanía de los placeres, configura, para Moore, un ilícito lógico, una falacia que llamaré «naturalista»; naturalista pues lo placentero es un ejemplo de un dato natural: un dato sensitivo, empíricamente corroborable.

Una manera eficaz de dismantelar la falacia naturalista es efectuando una pregunta. Lo que Moore, precisamente, llamará la «pregunta abierta» es un recurso reflexivo que permite abrir un «hueco lógico» en la operación de reducción de lo moral a lo no moral. Por caso, volviendo al ejemplo recién formulado, la pregunta abierta, oculta tras la operación falaz de reducción, podría ser esta: ¿pero es lo placentero, realmente, lo bueno moral?

La maniobra de Moore dejará una huella indeleble en la filosofía moral contemporánea. Inclusive, autores que se autodefinen como kantianos, en el sentido de verse como constructivistas morales, y no como realistas, plantean recursos similares al de Moore. Este es el caso de Christine Korsgaard (2000) para quien, lo que vayamos a considerar moralmente correcto es aquello que se presta a nuestro «asentimiento reflexivo». Lo que resulte moralmente correcto no puede reducirse, para la profesora americana, a aspectos naturales del mundo (empíricos o fácticos). Tan es así, que lo moralmente correcto depende, de últimas, no de nuestra identidad práctica empírica, sino de nuestra identidad práctica *necesaria*.

La *necesariedad* de la identidad práctica es un postulado —modal— muy exigente. Lo digo porque tal predicación implicaría que nuestra identidad —de por sí contingente—, pues podríamos tener otra, depende de hechos necesarios; hechos que denominé en mi lectura «anómala» de Kant más arriba, como hechos de la razón. Y al ser de la razón, los conceptos morales forjados en la fragua de la necesidad modal deberían, en última instancia, atravesar distintos mundos posibles; al menos, *ceteris paribus*, aquellos mundos donde ese agente humano conserve intactos los poderes racionales que consideramos válidos aquí.

Al introducir a Korsgaard aquí, solo he querido insertar un «símil parcial». Ello porque la autora americana asume, como principio epistémico de su sistema moral, una lectura constructivista estándar de Kant conteste con la cual, en tanto que agentes dotados de una racionalidad profunda o necesaria, somos «creadores» del reino de los fines morales; somos creadores, en *sum-*

*ma brevis*, de los principios y valores morales.<sup>3</sup> Repárese, a este respecto, que cuando hablé de Searle hice hincapié en la idea de «crear» y dije que volvería a este peculiar uso verbal. Pues bien, un realista moral no naturalista, platónico, tendría que asumir lo contrario: el mundo moral, constituido por las normas de lo bueno, lo justo, lo correcto no puede ser «creado» por agentes humanos. Si fuese creado, entonces se quebraría el supuesto de «independencia» entre dicho mundo y los agentes humanos que deben reconocerlo a fin de gobernarse por sus normas, valores y principios. Esto ubica el asunto de un realismo moral no naturalista en el dudoso ámbito de las categorías «*subspecie aeternis*». Es como si este mundo, al igual que el Dios cristiano, hubiese existido siempre, y existirá siempre, incluso si no existiésemos o dejásemos de existir. Esta forma de pensar hibridada con el realismo moral es, justamente, la que mayor escozor genera entre los filósofos morales. Las razones para tal repulsa son, cuando menos, estas tres:

- a) Solo un mundo deísta o teísta admitiría que la moral tenga ese origen en un mundo eterno increado por el hombre. El problema con este tipo de mundo es que la ética se comportaría como *ancilla theologiae*. Una ética así no solo carece de autonomía conceptual con respecto a la religión, sino que es ella misma, en sustancia, religión. Tal concepción de la ética retrocede mil años y nos instala en la ética medieval, olvidando los progresos de la ética moderna y contemporánea.
- b) Aunque cierta clase de independencia resulte crucial para entender la verdad objetiva de nuestras proposiciones morales básicas, no está claro que sea menester postular la absoluta inanidad del hombre. La intuición más seria es que el hombre, al menos, «coparticipa» de algún modo en la modelación de dicho mundo moral.
- c) El problema más serio, acaso, es el que ha denunciado Searle en su obra ya citada al comienzo del capítulo. Para Searle una pregunta fundamental es cómo conciliar la dimensión normativa (social, institucional, en su caso) con la dimensión natural. Esta última conforma los hechos más básicos sobre el mundo. El riesgo de un realismo moral no natural sería la generación de un mundo moral *singular*, sui generis. Esta singularidad parece defendible en que es altamente intuitivo que la normatividad de la moral debe preservarse de una reducción completa hacia lo natural. Sin embargo, la auto-

---

3 Téngase en cuenta que esta clase de creación no es «libre» de restricciones racionales en Kant; en cambio, en autores como Sartre (2003), de cuño existencialista, pareciera que el hombre crea valores en un sentido más radical. Aunque, no por esto, dicha creación no tiene alguna restricción; por ejemplo, para Sartre, el compromiso moral de un hombre con su acción individual compromete toda la humanidad. Así, esta formulación parece una combinación problemática de existencialismo y kantismo.

nomía de la moral no es absoluta. Si lo fuese, dejaríamos de explicar cómo se vincula lo causal con lo normativo, lo natural con lo moral. La moral no es unigénita. Estaría anclada en lo natural, a menos que generemos la extraña metafísica de los dos mundos: el natural y el moral. A estas «barbas metafísicas», parece que debiera aplicársele la navaja de Ockham.

Problemas como los recién enunciados parecen ser suficientes para dar pábulo al abandono, sin más, del realismo moral no naturalista. Sin embargo, antes de dar ese paso, creo que existe una alternativa más caritativa para repensar el realismo moral no naturalista.

La necesidad de encontrar una alternativa válida se asienta en una intuición, propia de los realistas morales, que considero plausible, *prima facie*: el mundo moral tiene carácter *normativo*. *Captar conceptualmente* de modo correcto este carácter parece requerir de preservar cierta autonomía para la ética respecto de lo natural. Sin embargo, la captación está incompleta si no damos cuenta de otra intuición que es bastante razonable: aquella según la cual, tiene que presuponerse una *distancia* entre el agente moral (el sujeto observador de las ciencias) y el mundo moral. Esto obedece a razones en principio lógicas. No hay observación, ni percepción alguna, si no admitimos la mentada distancia. Pero hay más que esto en juego: a la normatividad del mundo moral, a su, por lo menos, relativa autonomía se añade aquel presupuesto analítico implícito en el requisito de independencia ya aludido: me refiero a la necesidad de «objetividad» de las proposiciones morales; objetividad que aquí hago equivaler a la necesidad de preguntarnos si las mismas pueden ser verdaderas o falsas; más aún, parece que algunas tienen que ser necesariamente verdaderas, mientras que otras necesariamente falsas. Si ese no fuera el caso, ¿no habría mundo moral! ¿Cómo podría haberlo si, al menos, algunas proposiciones básicas con las que lo definimos, y lo identificamos, no fuesen conceptualmente verdaderas?

Todo lo dicho recientemente se dirige a defender otra alternativa para el realismo moral no natural que viene a ser la siguiente: el mundo moral, ciertamente no es «increado». Admítase por lo pronto esta premisa como una verdad axiomática. Sin embargo, de aquí no se sigue que los requisitos de normatividad, distancia y objetividad deban ser vulnerados.

Voy a sugerir, a continuación, que el mundo moral está *fuera de nosotros*. Pero no es menester que esta situación de externidad del mundo moral, respecto nuestro, implique que no coparticipamos, en modo alguno, en la modelación, definición, identificación y reconocimiento de dicho mundo moral. Aunque parezca paradójico, se podría sustentar la tesis de que dicho mun-

do está también «dentro» de nuestra mente moral. ¿Cómo es esto posible? ¿No se comete una contradicción al afirmar que ese mundo está afuera nuestro y adentro de nuestra mente? La contradicción, pienso, podría ser aparente. Ello porque ese mundo moral, junto a sus propiedades lógicas, podría pensarse como teniendo una elegante *estructura formal*.

Es formal porque ahora estoy pensando en la *forma* de las proposiciones morales y sus relaciones lógicas deductivas, o incluso de implicación pragmática. Esa estructura hace parte, a su vez, de la estructura de nuestra propia mente. La mente, y sus atributos, pueden pensarse perfectamente como formando una estructura. Si se me lo permite decir, esta es la intuición de los lingüistas como Noam Chomsky; intuición que cuenta, también, con evidencias empíricas. Tales evidencias son recogidas, con esfuerzo y paciencia, con avances y retrocesos, por los neuroéticos cuando intentan, como explica en esta misma obra González Lagier, calcar en papel la estructura naturalista del funcionamiento de la mente moral. La estructura de la mente, a su tiempo, debe verse reflejada, cuando menos parcialmente, en la estructura de nuestro lenguaje: nuestras proposiciones, sus propiedades, incluyendo el uso práctico de los predicados morales (bueno, correcto, justo, virtuoso, vicioso, etc.).

Si la versión alternativa que acabo de presentar tiene, por lo menos, algún grano de verdad, parece que hay aspectos altamente intuitivos en el realismo moral no natural. Más todavía: podríamos apelar a evidencias *empíricas* de los neuroéticos para hablar de la realidad natural del «cerebro» moral, el que está detrás, como fantasma o no, de la mente moral. O sea, hasta donde puede verse, la neuroética no tendría que ser enemiga de un realismo moral no natural.

Sin embargo, lo anterior no es lo que piensan autores como Martín Daguerre en esta misma obra. Para él, el realismo moral es conceptualmente falso. Ello es así si se admite, como él, que la única realidad moral es la de las operaciones del cerebro. Más que entidades ahí afuera, hay operaciones de la mente que son constatables a partir de un mundo natural.

El realismo moral naturalista puede asumir dos versiones. La primera, que quisiera llamar «cruda», es aquella según la cual lo moral se *reduce* a lo no moral (o natural). Spencer, Darwin, Bentham, podrían ser ejemplos de esta primera versión. Lo moral es aquello que se reduce, y que se explica, en la evolución, en la realidad biológica, o en la experimentación de sensaciones corporales como los placeres y la aversión al dolor.

La gran falla de San Andrés del realismo moral reduccionista que acabo de pintar es que, presumiblemente, *pace algunos de sus defensores*, cometería falacia naturalista. Por ejemplo, este es el caso, según creo, de Patricia Churchland (2020). Ella quiere explicar el fenómeno moral, en la obra *El cerebro*

*moral*, por el expediente de reducir el mismo a unidades básicas constituidas a partir del funcionamiento químico de la *oxitocina* en los organismos humanos. Esta hormona sería la «responsable» —la candidata a explicar— conductas como la cooperación. Sin embargo, pese a los dichos de Churchland, de que ella no cometería la falacia porque usa la idea de «derivación» en un sentido lógico «amplio», deja sin explicarse cómo tal amplitud inferencial sería lógicamente válida. Hasta donde veo, pasar del hecho de la oxitocina, a la justificación de valores como la cooperación, no está permitido por la lógica práctica. Ello si admitimos la ley de Hume ya mencionada en este capítulo.

A un realismo moral naturalista crudo se le añade, como mejora sustantiva, un realismo moral más «sofisticado». En mi opinión, cuando John Stuart Mill en su recordada obra *El utilitarismo*, distingue, criticando a Bentham, entre placeres bajos y elevados, procurando enriquecer un esquema basado exclusivamente en la mera *cantidad* de placeres y dolores, con el añadido del análisis de la *calidad* de los mismos, está formulando un realismo moral naturalista no crudo, o más sofisticado. Más todavía: cuando dos notables filósofos morales como Henry Sidgwick o George E. Moore se abroquelan tras un utilitarismo «ideal», parecen estar evidenciando la huella dejada por Mill. La idea de un placer elevado (como ejemplo de una experiencia moral) *no se reduce, o identifica*, con las sensaciones placenteras más básicas o crudas. En el fondo, Sidgwick, o Moore, están anticipando la idea que cobrará popularidad en la teoría moral, sobre todo a partir de su uso por parte de Richard Hare: la idea de «superveniencia».

El realismo moral naturalista más sofisticado no es un simple postular la moralidad *in rem*. No es que las categorías morales, como en *Metafísica*, sean la simple forma encarnada de la moral en los hechos empíricos o naturales. La noción mínima de superveniencia recién introducida tendrá éxito en la medida en que explique con sofisticada claridad el paso de lo no moral a lo moral.

Esto es, la superveniencia presupone que las propiedades morales supervienen, que no es lo mismo que reducir, a propiedades más básicas, crudas, empíricas o naturales. Si otros mundos satisfacen las leyes naturales que adoptamos en nuestro mundo natural, la superveniencia debería funcionar igualmente. La misma implica que si los hechos básicos son idénticos, inalterables, *ceteris paribus*, entonces las propiedades morales no pueden ser distintas. Mientras es una verdad conceptual que los hechos morales supervenidos tienen, como dicen los filósofos de la mente, «realizabilidad múltiple», es también una verdad que un mismo paquete de hechos naturales deberían dar lugar, en condiciones subjuntivas, al mismo paquete de hechos morales. Por la realizabilidad múltiple recién mencionada, la inversa no vale.

La idea de superveniencia permite, como dice Frank Jackson (1998) en su *From Metaphysics to Ethics* «localizar» un conjunto de propiedades morales (bueno, correcto, justo, valiente, bravo, envidioso, incorrecto, malo, injusto, cobarde, etc.).

La referida superveniencia no es un recurso exquisito solo del club de los filósofos morales. Los filósofos del arte, o los filósofos del color, utilizan el mismo recurso para pensar el modo lógico de reflexionar sobre el paso de lo natural a lo no natural. Más aún, en estas áreas de la filosofía recién nombradas, arte, color, también se da la gama de distinciones practicadas al comienzo de esta sección: entre realistas no naturalistas y realistas naturalistas crudos y más sofisticados.

Al esquema estructurado sobre la base de distinguir dos tipos de realismo moral: no natural y natural, alguien podría presionar algunas intuiciones (por ejemplo, la de independencia y distancia entre sujeto y mundo moral) para preguntar lo siguiente: ¿la realidad de las instituciones, concretamente de tipo moral, sería compatible con alguna forma inusitada de realismo moral? En adelante de cualquier eventual suspicacia de mi parte, contestaré a esta pregunta con un no. La dimensión institucional no es compatible con ninguna de las dos formas de realismo moral. Por tanto, *tertium non datur*.

Cuando repasamos las citas de Davidson, o de Searle, podríamos sentirnos confundidos con mi respuesta negativa. Pero la confusión, si estoy en lo cierto, no debería durar demasiado. Davidson, como Wittgenstein, Searle en modo similar, incluso Mackie, convergerían en que cierta clase de objetividad es posible y necesaria. Dicha objetividad vendría de contrastar una práctica lingüística compartida (Davidson, Wittgenstein) o una determinada realidad institucional (Searle y Mackie). Más aún, un naturalista, digamos estándar, podría admitir que la necesidad de objetividad de ciertos juicios normativos (morales, jurídicos, estéticos) no se pone en discusión. Sin embargo, ¿cuáles son los hacedores de verdad de nuestras proposiciones? Para estos filósofos, y me gustaría agregar a la lista a Hilary Putnam, los hacedores de verdad, que forman parte de una cierta clase de realidad con la cual contrastar el contenido de nuestros estados mentales y proposiciones lingüísticamente expresadas, no es «trascendente» en ninguno de los sentidos que postularía un realista moral. Más bien, son hacedores inmanentes a una práctica lingüística, a una realidad institucional o, para el caso de Putnam, a unos hechos que son de «este mundo».

Se puede salir muy rápido al ruedo para forzarme, al menos, a matizar mis aserciones. Por ejemplo, se me podría indicar que el realismo moral naturalista podría ser compatible con una cierta versión de la mente, del lenguaje, de nuestras prácticas e instituciones. Sin embargo, repárese en que los hechos



morales supervenientes en la versión naturalista del realismo moral deben guardar, todavía, una impresionante autonomía respecto de los hechos no morales que supervienen. De lo contrario, estaríamos, como arguye Michel Smith, eludiendo parte del «problema moral»; problema que debe consistir en poner en calma armónica dos ideas motrices: la normatividad de la moral, con la causalidad del mundo empírico. Justamente, la idea de superveniencia de los realistas naturalistas más sofisticados es un modo de encontrar esa calmada relación. Lo moral no puede reducirse (identificarse, si Usted prefiere) a lo no moral. ¿Qué significa este presumible misterio para nosotros, consistente en la duda de por qué hay que eludir la reducción de lo moral a lo no moral? ¿Qué *es* «lo moral»? Mi respuesta a esta última pregunta es que «lo moral», que está en juego en el análisis conceptual profundo, *no es lo convencional*. Por supuesto que «lo moral» en sentido positivo u ordinario se suele identificar con lo convencional. Ejemplo de ello podrían ser los sofistas que, como explica Barbara Cassin (2008), identificaban lo justo con las convenciones. Esta identificación podría ser «empíricamente» verdadera y daría rienda suelta a alguna forma de relativismo como aquella que Risieri Frondizi, con algo de sorna, llamaba la moral del «hombre adulto europeo detrás de una axiología objetivista» (2007:197). «Lo moral», en rigor del análisis, tiene que ver con una dimensión que incluso si de «base natural» no puede reducirse a esa base. El mundo moral puede ser uno con el mundo natural sin necesidad de reducción, tal como han probado Searle o aquellos que defienden el recurso a la superveniencia.

La necesidad de no reducir —o identificar— lo moral con lo convencional ha sido entrevista por autores como Herbert Hart o John Mackie, para citar dos casos conspicuos, cuando han remarcado que las convenciones de la moral ordinaria, pueden ser erróneas, crueles, irracionales, perversas. Alguna vez, la esclavitud fue también una «institución» y satisfacía aquel tipo de norma que Searle llama «constitutiva» según la cual «x cuenta como Y en las circunstancias c». En contraposición con «lo moral», las instituciones (los clubes, las iglesias de diversos credos, los estados, las familias, etc.) pueden ser, como arguye por ejemplo Leo Strauss (2015) opresoras, irracionales, crueles, perversas.

Precisamente, el dato que se acaba de consignar con respecto a la falibilidad de las instituciones (su eventual irracionalidad o perversión) explica que no siempre los hechos morales, entendidos en sentido crítico, puedan alinearse lógicamente a los hechos institucionales *searleanos*. Más aún, la invocación a hechos morales en sentido estricto o crítico, sean estos concebidos como sustancias platónicas, o como hechos supervenientes, puede hacer que la dimensión institucional resulte objetable. Por ejemplo, este puede ser el

caso de los hechos institucionales que configuran un subconjunto normativo como el derecho positivo de un país. El derecho puede ser «nazi», una sociedad puede experimentar el *apartheid*, o incluso, siendo un derecho justo en términos generales, puede haber normas «injustas» desde el punto de vista de su aplicación particular. La no alineación entre hecho moral e institucional puede revelarse, asimismo, por signos que van de la simple protesta callejera por una política impopular, hasta los actos más directos de desobediencia civil o resistencia a la autoridad. Con lo cual, la predicación de hechos morales, efectuada por el realista moral, presenta una doble cara: quizás una ventaja en términos de permitir la corrección de ciertas prácticas institucionales como, por caso, el derecho positivo de un país. Pero, a pie seguido, surge la perturbadora dificultad de que no podamos, a veces, explicar cómo el razonamiento práctico —*verbigratia*, moral— puede estar alineado, en una unidad lógica armónica, con los hechos institucionales. Más todavía, los hechos morales pueden erosionar las bases de lo que, por ejemplo, un positivista jurídico, llamaría la «autoridad» del Derecho (Iosa, 2017).

Es completamente irrefutable, sin embargo, el hecho de que las instituciones pueden funcionar como hacedores de verdad de ciertas proposiciones. Un ejemplo nítido de ello es el Derecho. Los juristas suelen expresar proposiciones, por ejemplo, el juez J tiene competencia en el caso x, solo si C. Esta clase de proposiciones son verdaderas o falsas según cómo sean, «realmente», las instituciones jurídicas. Hay realismo en la dimensión institucional. Hay «objetividad» de las proposiciones e, inclusive, hay cierta distancia entre esas instituciones y sus creadores. Pero «lo moral», estrictamente comprendido el sintagma, no es eso, no puede ser eso, no puede ser «meramente creado», o «meramente proyectado» como sostiene un antirrealista (o cuasirrealista) como Simon Blackburn de acuerdo con la explicación en esta misma obra de Federico Arena.

Lo moral, en sentido *crítico*, no es convencional. Ello no significa, que no podamos, por hipótesis, en cierto tiempo T<sub>1</sub>, llegar a ciertos acuerdos morales profundos y, eventualmente, estables.

Por si acaso mi respuesta no fuese convincente voy a incluir, en lo que sigue, un argumento explicitado por George E. Moore (1997) en su *Principia Ethica*. Es el argumento del «valor intrínseco». A mi entender, con enorme perspicacia filosófica, el antecesor de Wittgenstein en Cambridge, sostiene que la ética, bien entendida, se vincula no tanto con la noción de «objetividad moral», como con el concepto de valor intrínseco. Y creo que tiene algo de razón. Por lo que he mostrado, las instituciones, o las convenciones morales si usted quiere, pueden garantizar cierta objetividad al funcionar como hacedores de verdad de ciertas proposiciones. Empero, es muy importante, desde ya, advertir que son hacedores de verdad, o funciones lógicas de ver-

dad, «relativos» a un conjunto de instituciones o convenciones que, a su vez, tienen *historia* y, son, por esto, *contingentes*. No es posible ser un universalista moral estricto desde una visión puramente convencional de la moral. Desde la moral positiva, como sugiere en esta obra Graciela Vidiella, no es sencillo llevar agua para el molino del universalismo moral *stricto sensu*.

Los valores, los morales por caso, pueden ser relativos o no relativos. Los relativos, por lo general, suponen una racionalidad instrumental: los valores son relativos a otros valores en el sentido en que hay medios que son relativos a fines o metas. Volviendo al ejemplo del placer. Supóngase que, al mejor estilo Sade, a una persona le da mucho placer atormentar a las mujeres sin apelar a su consentimiento, y emplea instrumentos como púas, alfileres o cosas por el estilo al tener sexo con ellas. Frente a esta tendencia o impulso se podría pergeñar un silogismo instrumentalista como este: el sujeto *s* quiere tener placeres sádicos con las mujeres con las que se acuesta (valor final de su acción); utilizar púas o alfileres para herirlas es el medio adecuado; ergo, *s* «debe» (en sentido hipotético) utilizar púas o alfileres para herirlas.

Frente a un ejemplo mental como el de recién, lo menos que se espera de un agente moral verdadero es que *se pregunte* si su acción y su razonamiento son *moralmente* correctos. Incluso, si no se hiciese la pregunta, la misma «estaría ahí abierta», disponible. Justamente, esta es una fuerte intuición del realista moral: aquella según la cual hay preguntas que están abiertas, *lo veamos o no, lo queramos o no*.

Por lo pronto, digamos que también hay valores no relativos; valores que, en su *Conferencia de ética*, Wittgenstein (2005) vio como definitorios de lo moral y, al mismo, tiempo, como imposibles dentro de un lenguaje con «límites».

Estos valores no relativos, son equivalentes al tipo de valores «intrínsecos» mentados por Moore y examinados en el capítulo escrito por Facundo García Valverde. Son valores de esta clase aquellos que valen por sí mismos, no como medios para otros fines. Tales valores, para Moore, eran relativamente independientes de los hechos naturales. Cuando nuestro filósofo estaba seducido por la metafísica de tipo mereológico enfocada en la relación peculiar entre partes y de las partes con un todo, los valores intrínsecos eran «algo más» que la mera reunión o colección de pedazos. Es por esta razón conceptual que el valor moral, que en esencia es intrínseco, no se reduce, no se explica, como una derivación de partes reunidas. Si una de esas partes es la experimentación de cierto tipo de placer, ello no hace que ese placer tenga valor intrínseco por su sola presencia. Hay que preguntarse qué clase de cosas son valiosas *per se, por sí mismas*. En definitiva, «lo moral» se reenvía al reino de los axiomas, de lo que de suyo es valioso, y sobre la base de lo cual derivamos luego predicados morales. Así, el realista moral, como Moore dice, debería pegarse

a la noción de valor intrínseco, más que obsesionarse con la mera objetividad de las proposiciones o juicios morales.

## **Lo cognitivo y lo práctico**

Deseo finalizar este capítulo, considerando un tema pendiente en el tratamiento del realismo moral como doctrina metaética. Me refiero concretamente a la naturaleza del contenido de las proposiciones o juicios morales. Una proposición o juicio, suele tener formas y contenidos como los siguientes: «el aborto de un embrión con tres semanas de vida intrauterina está permitido si la gestante ha sido violada»; «en caso de una enfermedad terminal, dolorosa e irreversible, la eutanasia activa es moralmente permisible», «las guerras de agresión, basadas en motivos falsos, como fue el caso de la invasión que George Bush hijo decidió practicar contra Irak, son injustas», etc. La forma lógica de las proposiciones morales está determinada por los operadores deónticos (obligatorio, justo, correcto, prohibido, permitido, incorrecto, injusto, bueno, malo, etc.) o aretaicos (valiente, cobarde, noble o envidioso, etc.). El «contenido» está dado por afirmaciones o negaciones como las enunciadas con los ejemplos precedentes.

Ahora bien, suele endilgársele al realista moral, cualquiera sea su clase, una caracterización «cognitiva» de la moralidad que contrastaría, a la postre, con caracterizaciones «no cognitivas» de la moral. Richard Hare (1997) no vio la conveniencia de clasificar así a las teorías éticas y prefirió, como sabemos, distinguir entre teorías descriptivas y no descriptivas (o prescriptivas). Los realistas morales, para él, eran descriptivistas, mientras que él, a diferencia de ellos, era un prescriptivista.

En añadido de las dos distinciones arriba delineadas (cognitivo vs no cognitivo, descriptivo vs prescriptivo), autores como, por ejemplo, Christine Korsgaard, han indicado que el gran inconveniente del realismo moral estriba en su carácter epistémico y en su falta de carácter «práctico». O sea, para la filósofa nombrada, los realistas morales solo estarían obsesionados por el carácter «cognitivo» (o descriptivo de Hare) de las proposiciones morales. Solo así sería capturada la intuición de que, al menos, algunas proposiciones básicas de la moral pueden y deber ser verdaderas, mientras que otras falsas. Sin embargo, la crítica de Korsgaard a la presunta obsesión realista en lo epistémico, en detrimento de lo práctico, es que la ética tiene una vocación eminentemente normativa. Esta vocación se traduce en que la ética —que llamé al comienzo sustantiva o normativa— establece las pautas, esto es, las normas, principios y valores a los que deberíamos ajustarnos o que deberían motivarnos.

Es claro, por lo pronto, que verdad y carácter práctico son atributos analíticamente diferenciables. El punto máximo de la crítica al realismo sería que, por centrarse en el carácter cognitivo, pierde de vista cómo explicar la motivación moral.

Pienso que esta clase de objeciones es injusta con el realismo moral. Hacer foco en el carácter cognitivo de la moral, no implica *negar su carácter práctico*. Pues, debería ser de suyo evidente, que «hacer foco», o «prestar atención», no es lo mismo que negar o demeritar la importancia del carácter práctico de la moral. Cualquier ético admitiría, según creo, que la moralidad tiene precisamente en su normatividad, y en su aspiración a la practicalidad, buena parte de su esencia. El paso lógico independiente, sin embargo, es preguntarse por la posibilidad o no de verdad de, al menos, algunos juicios morales. Una preocupación, si se quiere, que bordea temas centrales para cualquier filósofo como son el significado, la verdad y la objetividad de los términos, en este caso morales, de que disponemos.

Hay un punto final que quiero hacer, en conexión mediata con lo anterior, que me parece característicamente eludido por muchos filósofos morales. Me refiero a la clasificación misma de las teorías éticas entre cognitivistas y no cognitivistas, descriptivistas y prescriptivistas, de la que partí al comienzo de esta sección.

Estoy convencido, parcialmente, de que es una clasificación engañosa. Mis razones son éstas:

a) Autores clasificados no cognitivistas como, por ejemplo, Allan Gibbard o Richard Hare, son «no cognitivistas» en un sentido un tanto confuso. Gibbard, *qua* neo-expresivista, Hare, como prescriptivista, admiten en sus sistemas morales rasgos de racionalidad, o constreñimientos racionales al juicio moral, cualesquiera este sea, muy fuertes y evidentes. En el caso del primero, se postula la necesidad de una «aceptación» del sistema moral en virtud del cual se articulan luego las actitudes de aprobación o rechazo. Dado que este sistema procura la cooperación, no es cierto que un expresivista se desentienda de constreñimientos racionales. Vale, *mutatis mutandis*, lo dicho en relación con Gibbard, para Hare. El defensor del prescriptivismo, asienta lógicamente el «néustico» (la prescripción moral) en el frástico (los hechos no morales). Recuérdese que señalé, páginas atrás, que Hare hizo explícita la superveniencia, la que va del frástico al néustico. Este néustico además, debe ser susceptible de satisfacer la ley del peritropo: o sea de ponerse en los zapatos de los otros, cada vez que emitimos juicios morales. En buen castellano, las prescripciones de Hare tienen que aspirar a la universalidad en los casos ordinarios, o sea, los poco problemáticos; en casos recalcitrantes, será necesaria articular razonamientos consecuencialistas,

- con lo cual, creo, habría leyes empíricas tácitas que deben ser satisfechas al razonar en forma consecuencial.
- b) Si lo anterior es una descripción sumaria básicamente correcta, es hora de repensar el empleo de términos como cognitivo, no cognitivo, descriptivo y prescriptivo para clasificar teorías éticas. Aunque los autores expresivistas o prescriptivistas, como los recién comentados, no hablen de verdad, no es cierto que despojen al juicio moral de ciertos rasgos de racionalidad empírica y práctica. Con lo cual, aunque se privilegie términos como no cognitivistas y prescriptivistas, no debería ocultarse el importante hecho de que estos autores no creen que los juicios morales sean subjetivos «a secas» como era dable pensar en el viejo expresivismo de Ayer, o el emotivismo de Stevenson. Los juicios morales de rasgos acusadamente prácticos deben referirse a sistemas normativos (que funcionan con operadores lógicos), o bien el néustico, las prescripciones, no son un episodio de salto arbitrario de lo no moral hacia lo moral.
- c) Finalmente, la fruta del postre que faltaba. Dije, en un tramo de mi argumentación que una intuición poderosa del realismo moral está ligada a la idea según la cual hay ciertas cosas que debemos hacer, o debemos abstenernos de hacer, *lo queramos o no*. Con franqueza, quizás sea esta la parte más atrayente del realismo moral, pues considero es su intuición más válida. En el fondo, es la idea que late en un filósofo —considerado como antirrealista, constructivista o realista anómalo— como Immanuel Kant. Lo moral, estrictamente, puede, y en ocasiones debe, oponerse a las inclinaciones.

La idea según la cual hay cosas morales que están ahí, lo queramos o no, hace que los hechos morales se asemejen, en un punto, a la existencia de montañas o ríos, cuando decimos que están ahí, los veamos o no, queramos verlos o no. Decir esto, no es para un realista moral un «desliz de la pluma filosófica», sino que reafirma una intuición extendida, pero que debe seguir siendo clarificada. Una pregunta inquietante —todavía pendiente de clarificación— sería si estos hechos morales son el punto de hervor «justo» de la objetividad moral.

Aristóteles planteaba la idea de hecho moral, me parece, en términos que hoy denominaríamos de «absolutos morales», es decir, conductas que, como las obligaciones perfectas negativas de Kant, deben ser cumplidas por los agentes morales, para ser considerados morales.

En el trasfondo de lo dicho, el realista moral, *pace* la crítica de filósofos como Korsgaard, es un realista que yo denominaría «normativo». Una correcta comprensión de lo moral, *rectius*, debería ser capaz de atrapar este elemento de «autoridad» de lo moral (de aquí el empleo reciente del predicado

«normativo»); autoridad que descansa en la intuición de que hay cosas que «se nos escapan»; cosas, o hechos morales, con respecto a los cuales, aunque suene raro decirlo, somos *un tanto pasivos*.<sup>4</sup> Esta pasividad equivale a aceptar esa realidad moral normativa, incluso si desmiente ciertas creencias, o contraría ciertos deseos o voliciones. En castellano, esto equivale, lógicamente, a admitir que la moral está ahí afuera, en algún sentido metafórico pero fructífero, quizás, de la palabra; no podemos «disponer de ella» como queramos, tal como transmite David Enoch (2011) en su *Taking morality seriously*; un título que, curiosamente, un realista moral utiliza parafraseando el título de otra obra, la del «antirrealista», pero objetivista de todos modos, Ronald Dworkin, en su *Taking rights seriously*.

Como acabo de indicar, mi reconstrucción se hace extensible a lo que David Enoch llama la «indispensabilidad» de la moral. Negar que hay hechos morales necesarios, para él implica caer en alguna suerte de contradicción lógica pues queremos dar cuenta de lo moral, borrando su esencia al mismo tiempo. Su esencia, es su normatividad, entendida ahora como autoridad sobre nosotros. Claro que, en algún sentido, esa moral necesita «ser nuestra», ser reconocida y motivar. Y en algún sentido, como he explicado páginas atrás, ya «es» nuestra, producto de la estructura formal de nuestra mente moral.

Sin embargo, para el realista moral, si es que he logrado exponerlo aquí con claridad —y por qué no, de vivacidad— no somos «creadores» a secas, o proyectores a secas de propiedades morales, como sostendrían (algunos) constructivistas, ficcionalistas, expresivistas, o cuasirrealistas morales.<sup>5</sup> Las propiedades morales que atribuimos a diversos aspectos del cosmos guardan, pese a todo genio creador, una sagrada independencia y esto es lo que he intentado captar en este capítulo.

---

4 Piénsese, por ejemplo, en la idea extendida que una filósofa como Simone Weil (2007) tiene sobre el rendirse pasivamente a lo que resulta moralmente sagrado.

5 Digo «algunos» porque no se me escapa que hay variantes dentro del amplio arco de posturas metaéticas convencionalistas, expresivistas, etc., que intentan mostrar que existe una «realidad moral» (independiente en algún sentido) sin recaer, por esto, en un compromiso metafísico denso o plúmbeo. Así, por ejemplo, es el caso de Lovibond (1993) o de la wittgensteiniana Sol Yuan en este mismo libro.

## Coda

Hasta aquí he querido ser un «buen abogado» de las causas del realismo moral. Sin embargo, hecha su presentación casi apologética, queda, como resto, unos apurados y fragmentarios apuntes de sus dificultades.

La primera dificultad está en el corazón de la palabra «metaética», una palabra sobrecargada de una cierta manera de entender —bastante esquemática— a la filosofía analítica. Conforme este esquematismo, el filósofo metaético no hace implicaciones sustantivas para la vida moral de ningún tipo. Hasta aquí se podría decir que esto es inteligible. Pero desafortunadamente es muy insuficiente, inclusive lo es para una metaética que desee comprender —de verdad— la compleja riqueza de la vida moral o de las prácticas morales.

La segunda dificultad, más cercana a la letra ya, que, al espíritu de los problemas, como la anterior, es que la postulación de una realidad moral *eidética*, aun si intenta noblemente asegurar la necesidad de objetividad moral en un sentido fuerte, pierde demasiado de vista. Unos hechos morales independientes de mis creencias, y deseos, no solo admiten una realidad muy remota con capacidad intrigante para motivarme a actuar moralmente, sino que presupone algo mucho más problemático. Si el filósofo sale de su ingenuidad, y es honesto hasta mancharse, sabrá entender que una realidad moral respecto de la cual los agentes morales están *exiliados*, una realidad que no pueden hacer propia, es una realidad fantástica. El modo más efectivo que un agente tiene de hacer propio un reino moral objetivo es por medio de lo que ya Aristóteles intuyó: la experiencia moral. Es allí donde el agente debe asumir el punto de vista subjetivo, de la primera persona, sobre lo que percibe, siente, experimenta. Solo una metaética «analítica» en el sentido esquemático descrito, puede dejar de lado la relevancia crucial, por derecho propio, de la experiencia moral. Más cualificadamente hablando, la experiencia, de digamos, el dolor, la enfermedad, la del soldado que regresa de una guerra, del hombre atribulado por sus deudas, etc., son fuentes morales en un sentido profundo: se sitúan en el espacio propio, el del agente *que asume un punto de vista* y que está arrojado —o condenado— a decidir de un modo u otro. El papel de las emociones en la toma de esas decisiones, por ejemplo, de la angustia, la culpa, la tristeza, el miedo, el amor, no están retratados por la metaética obsesiva del realismo moral como lugarteniente de la objetividad. La analítica de las emociones suele formar parte de un cajón de sastre separado del cajón en el que entra la realidad moral.

Desde un punto de vista analítico amplio, la experiencia moral podría ser, inclusive, uno de los «hechos morales» más básicos para la teoría moral. Empero, no se trataría de un hecho moral meramente «objetivo», tal como re-



zaría el eslogan de un realista moral estándar. No meramente porque no es un hecho en algún sentido «ajeno» a nosotros, a nuestras creencias y/o emociones morales. Sino que la experiencia moral sería otro nombre para hablar de *nuestro* «punto de vista» sobre los temas de relevancia moral, incluso de aquello que, para nosotros, contaría como algo de color moral. Desafortunadamente, la expresión «hechos morales» en manos de teóricos morales suele reenviarse —en forma preponderante o excluyente— a una noción basada en la *tercera persona*. Y esto por definición de hecho moral objetivo: aquel que es independiente de nuestros puntos de vista. Pero, si estoy en lo cierto, esta idea de hecho no captura un dato esencial: *nuestro* lugar en el mundo moral. Y, así, es como si estuviésemos exiliados de nosotros mismos, tan afanosos de hallar el punto de vista del universo en materia moral.

Al de por sí complicado tema de la primera persona (Finkelstein, 2010; Wilson, 2016), se añade el de la segunda. Compartimos un mundo o, por lo menos, se supone que en algunos puntos lo hacemos y esto, por supuesto, es compatible con los desacuerdos. Compartir un mundo, o participar de una práctica, parece que deja lucir un aporte de la tradición hermenéutica que un filósofo analítico no debiera desdeñar. El aporte básico que tengo en mente se conecta con investigar el grado de significación que tenga —a la hora de identificar un hecho moral— las maneras en que interpretamos el mundo, los modos en que, juntos, armónicamente o no, atribuimos valor y significado a las cosas (por ejemplo, Ricoeur, 2006). Ambos ítems mentales, primera y segunda persona, entonces, demandan de los filósofos confiados, algo soberbiamente a veces, en la existencia de una realidad trazada por hechos morales objetivos, de una mirada más compleja. Esos hechos no pueden como sostuve más arriba, exiliar nuestros puntos de vista. Y nuestros puntos de vista a veces se entrecruzan con los de otros agentes con los que, en algún sentido, y en algunos puntos compartimos un mundo. Fue Donald Davidson quien llamó a una compilación de trabajos suyos con la sugestiva tríada *Subjetivo, intersubjetivo, objetivo*. Son los componentes que, como filósofos morales, tenemos que seguir pensando, en términos atómicos, y en términos holísticos. Y pensar estos componentes, quizás, exija de nosotros trascender consideraciones enmarcadas solamente en la así llamada «tradición analítica». No hay en juego solamente aspectos mentales, o semánticos de estos componentes. Y hay tradiciones, como la hermenéutica, inclusive la existencialista, para poner solamente dos ejemplos, que probablemente nos ayuden a ver mejor la sustancia de las relaciones entre los tres componentes. La realidad moral, entonces, debería ser la síntesis de los tres ítems.

Reitero: dar cuenta del rico tapiz mental constituido por la primera, segunda y tercera persona, tapiz vinculado, quizás, al ámbito de comprensión her-

menéutica, e incluso, y también, quizás, existencial, debe hacerse tomando en serio la noción de experiencia moral. Mi intuición es que, sin la mediación de esta noción, la idea, —sobre todo, metaética— de realidad moral es demasiado abstracta y, a lo mejor, peligrosamente vacía.

## Referencias bibliográficas

- Arendt, H. (2001). *El concepto de amor en San Agustín* (trad. de Agustín Serrano de Haro). Encuentro Ediciones.
- Butler, J. (2006). *The Works of Bishop Butler*. Boydell & Brewer.
- Cassin, B. (2008). *El efecto sofístico* (trad. de Horacio Pons). Fondo de Cultura Económica.
- Churchland, P. (2020). *El cerebro moral. Lo que la neurociencia nos cuenta sobre la moralidad* (trad. de Carme Font Paz y Montserrat A. Fernández). Paidós.
- Davidson, D. (2003). *Subjetivo, intersubjetivo, objetivo* (trad. de Olga Fernández Prad). Cátedra.
- Dworkin, R. (2014). *Justicia para erizos* (trad. de Horacio Pons). Fondo de Cultura Económica.
- De Libera, A. (2015). *La cuestión de los universales. De Platón a fines de la edad media* (trad. de Marité Miccio). Prometeo.
- Enoch, D. (2011). *Taking morality seriously: A defense of robust realism*. Oxford University Press.
- Ferraris, M. (2012). *Manifiesto del nuevo realismo* (trad. de José Blanco Jiménez). Ariadna ediciones.
- Finkelstein, D. (2010). *La expresión y lo interno* (trad. de Lino San Juan). KRK ediciones.
- Frondisi, R. (2007). *¿Qué son los valores?* Fondo de Cultura Económica.
- Hare, R. (1997). *Ordenando la ética. Una clasificación de las teorías éticas* (trad. de Joan Vergés Gifra). Ariel.
- Iosa, J. (2017). *El conflicto entre autoridad y autonomía*. Fontamara.
- Jackson, F. (1998). *From metaphysics to ethics: a defence of conceptual analysis*. Oxford University Press.
- Korsgaard, C. (2000). *Las fuentes de la normatividad* (trad. de Laura Lecuona y Laura Manríquez; rev. de Fabiola Rivera). Instituto de Investigaciones Filosóficas (UNAM).
- Lafont, C. (2002). Realismo y constructivismo en la teoría moral kantiana: el ejemplo de la ética del discurso. *Isegoría*, (27), 115–129.
- Lovibond, S. (1993). *Realism and imagination in ethics*. University of Minnesota Press.
- Mackie, J. (2000). *Ética. La invención de lo bueno y lo malo* (trad. de Tomás F. Auz). Gedisa.
- Moore, G.E. (1997). *Principia ethica* (trad. de Adolfo G. Díaz e Isabel Stellino). Instituto de Investigaciones Filosóficas (UNAM).
- Moreso, J. (2010). Tomates, hongos y significado jurídico. En J.J. Moreso, L. Prieto Sanchís y J. Ferrer Beltrán (Eds.), *Los desacuerdos en el Derecho* (pp. 1548). Fundación Coloquio Jurídico Europeo.
- Parfit, D. (2011). *On what matters*. Vol. 2. Oxford University Press.

- Ricoeur, P. (2006). *Sí mismo como otro*. Siglo XXI.
- Sartre, J.P. (2003). *El existencialismo es un humanismo* (trad. de Manuel Lamana). Losada.
- Searle, J. (2017). *Creando el mundo social. La estructura de la civilización humana* (trad. de Juan Bostelmann). Paidós.
- Strauss, L. (2015). *Derecho natural e historia* (trad. de Luciano Nosetto; rev. de Dolores Amat). Prometeo.
- Weil, S. (2007). *La gravedad y la gracia* (trad. de Carlos Ortega). Trotta.
- Wilson, C. (2016). *Metaethics from a first personal standpoint. An introduction to moral philosophy*. Open Book Publishers.
- Wittgenstein, L. (2005). *Una conferencia de ética* (trad. de Alejandro Tomasini Bassols). Instituto de Investigaciones Filosóficas (UNAM).

## 4. Antirrealismo y Cuasirrealismo moral

*Federico José Arena*

Estamos acostumbrados a usar el lenguaje para describir cómo es el mundo que nos rodea y es muy poco frecuente que ello nos parezca misterioso. Cuando decimos «Hoy el cielo está despejado» o «Hay pocos vehículos circulando por la calle» sabemos, más allá de algunos casos complejos, que tales enunciados tienen la pretensión de describir un cierto estado de cosas y también somos capaces de determinar si logran realizar ese objetivo. Es decir, somos capaces de determinar si son verdaderos o falsos. Por ello, en caso de que desacordemos acerca del valor de verdad de un enunciado de este tipo, sabemos dónde mirar y qué buscar para poner fin al desacuerdo. Es decir, somos capaces de realizar una investigación empírica con ese objetivo.

También estamos acostumbrados a usar el lenguaje para llevar a cabo valoraciones. O, más precisamente, para indicar el estatus moral de cierta acción o formular normas morales. Enunciados como «La interrupción voluntaria del embarazo es una libre decisión de la mujer sobre su propio cuerpo» o «La tortura de personas inocentes es un acto aberrante» son de este tipo. Si bien en ciertos casos consideramos que estos enunciados funcionan igual que los anteriores, en el sentido que algunos son verdaderos y otros falsos, con frecuencia resulta un misterio establecer de qué modo determinar su valor de verdad. Es decir, no resulta sencillo determinar dónde mirar para establecer la verdad o falsedad de un enunciado moral. Del mismo modo, tampoco resulta sencillo, frente a un desacuerdo, establecer quién formula enunciados morales verdaderos y quién no, más allá de que estemos convencidos de que el desacuerdo puede resolverse.

Bien, el Realismo moral es, en general, la posición según la cual los enunciados morales funcionan tal como lo hace el primer tipo de enunciados. Es decir, según el Realismo moral los enunciados morales son acerca de un estado de cosas y, por lo tanto, pueden ser verdaderos o falsos según coincidan o no con el estado de cosas referido.<sup>1</sup> Por ello, una de las tareas principales que debe enfrentar quien defienda al Realismo moral es precisar cuál es ese estado de cosas (o hechos morales/normativos) del cual depende la verdad o

---

1 En este capítulo se mencionarán posiciones o teorías que están desarrolladas con mejor y más detalle en otras contribuciones de este libro, por lo que remito al lector a esos textos para encontrar allí matices y distinciones ulteriores. De todos modos, no podré evitar introducir algunas referencias propias, a los simples fines de una mínima completitud del presente capítulo.

falsedad de los enunciados morales.<sup>2</sup> Por el contrario, el conjunto de teorías acerca del discurso moral que analizaré en este texto, englobadas bajo el rótulo «Antirrealismo moral» sostienen que el misterio se debe a una falta de comprensión del funcionamiento de ese discurso. Creer que existen enunciados morales verdaderos en cuanto describen adecuadamente hechos morales (o normativos) es una equivocación. Una de las razones de la equivocación es que no hay tales hechos morales (o normativos). Como veremos, hasta aquí llega el acuerdo dentro del Antirrealismo, pues es posible distinguir al menos tres variantes: el no-cognitivismo, la teoría del error y el cuasirrealismo. Todas ellas convergen en negar la tesis según la cual existen hechos morales (o normativos). Pero discrepan acerca de cómo ello impacta en el valor de verdad de los enunciados morales y en su funcionamiento.

En lo que sigue, precisaré la diferencia entre Realismo y Antirrealismo acerca del discurso moral. Luego introduciré las diferentes variantes del Antirrealismo, comenzando por el no cognitivismo. Al respecto, propondré tanto un argumento en su defensa como una crítica. A continuación, mencionaré los modos en que las dos variantes siguientes, del error y el Cuasirrealismo, pretenden responder a esa crítica y respecto de cada una de ellas ofreceré también un argumento a su favor. Concluiré luego con un argumento general en contra del Antirrealismo. Obviamente esta exposición no pretende agotar, ni mucho menos, los puntos en discusión.

## **La diferencia entre Realismo y Antirrealismo sobre el discurso moral**

Con cierta frecuencia se leen u oyen expresiones del tipo: «Bueno, esa es tu opinión, yo tengo la mía», o también «Que cada uno piense lo que quiera», entre otras varias del género. Estas expresiones suelen producirse a menudo en el contexto de discusiones sobre lo que es debido hacer, es decir, cuando las personas involucradas desacuerdan acerca de qué respuesta dar a interrogantes morales: si está justificado que el Estado cobre impuestos a los que más tienen para mejorar la situación de los más pobres; si debería usarse o no el lenguaje inclusivo en textos oficiales; si las mujeres deberían o no tener siempre el derecho a decidir sobre su propio cuerpo, y un largo etcétera. Dada la relevancia de estas cuestiones, resulta extraño que las personas involucradas,

---

2 Con algunas excepciones, claro. Ronald Dworkin, por ejemplo, sostiene que la verdad o falsedad de los enunciados morales no depende de que coincidan con cierto estado de cosas o hechos morales, pues no hay tal cosa, sino que depende de que se encuentren apoyados en un buen argumento moral. Véase la defensa de esta posición en Dworkin (2014:57–59).

acuerden terminar la discusión con expresiones como las que señalé al principio. La extrañeza se debe a que esas expresiones dejan traslucir una especie de relativismo acerca de las cuestiones morales.<sup>3</sup> Y, a pesar de la popularidad de este tipo de expresiones filo-relativistas, buena parte de las personas que las pronuncian no estarían de acuerdo con las consecuencias que se seguirían de un relativismo global. Es decir, no estarían de acuerdo con un relativismo global según el cual ninguna cuestión moral tiene una única respuesta correcta. Muchas personas filo-relativistas no estarían dispuestas a terminar de ese modo discusiones acerca del homicidio, la violación o incluso la interrupción voluntaria del embarazo, por ejemplo. Acerca de estos temas la defensa de la propia posición suele ser más enfática, en el sentido de que quien afirma lo contrario está equivocado.

El Realismo moral es la posición que mejor parece encajar con este último modo de percibir el discurso moral. El realismo, en general, es el punto de vista que atribuye valor de verdad a los enunciados proferidos dentro de un determinado discurso puesto que, afirma, existen las entidades que hacen verdaderos o falsos los enunciados proferidos dentro del discurso en cuestión. Sobre esta base, respecto del discurso moral podemos definir al Realismo moral a partir de las dos tesis siguientes:

*(1) Realismo moral*

- (i) Tesis semántica: los enunciados morales son verdaderos o falsos y algunos son verdaderos (tesis del cognitivismo semántico).
- (ii) Tesis ontológica: existen hechos morales o normativos (de los que depende la verdad o falsedad de los enunciados morales).

El Antirrealismo, en cambio, sostiene la posición opuesta. Así, puede ser definido a partir de las dos tesis siguientes:

*(2) Antirrealismo moral*

- (i) Tesis semántica: no hay enunciados morales verdaderos.
- (ii) Tesis ontológica: no existen hechos normativos.

---

<sup>3</sup> Por relativismo entiendo la idea de que dada la diversidad de puntos de vista acerca de determinada cuestión, no hay una única respuesta correcta. Suele distinguirse entre relativismo descriptivo y relativismo normativo. El primero se caracteriza por la tesis empírica según la cual diferentes culturas consideran verdaderos diferentes enunciados morales. El segundo se caracteriza por la tesis normativa según la cual la verdad o falsedad de un enunciado moral depende de la propia cultura. Véase Harman y Thomson (1996).

Un primer punto a señalar es que el Antirrealismo no es un punto de vista relativista, pues no sostiene que el valor de verdad de los enunciados morales depende de lo que piense quien lo enuncia, sino que, por el contrario, afirma que no hay enunciados morales verdaderos.

Ahora bien, como se habrá ya advertido, la tesis (i) del Antirrealismo es ambigua, dado que puede ser interpretada de dos modos diferentes. Por un lado, que no haya enunciados morales verdaderos puede deberse a que, en realidad, los enunciados morales carecen de valor de verdad (es decir, no son ni verdaderos ni falsos). Por otro lado, que no haya enunciados morales verdaderos puede deberse a que, en realidad, los enunciados morales son todos, inevitablemente, falsos. Esta ambigüedad es la que produce la división entre el Antirrealismo no cognitivista (los enunciados morales carecen de valor de verdad) y el Antirrealismo del error (los enunciados morales son todos, inevitablemente, falsos).

A partir de esta primera caracterización de las dos variantes del Antirrealismo, como simple negación de las tesis del Realismo moral, avanzaré ahora en la presentación de sus tesis positivas centrales.

### **Antirrealismo no cognitivista**

Además de usar el lenguaje para describir estados de cosas, estamos también acostumbrados a usar el lenguaje para expresar nuestros estados de ánimo, emociones o actitudes. Cuando decimos durante un partido de fútbol «¡Vamos Argentina!», o cuando decimos, al ver que ha rebalsado el agua de la bañera, «¡Qué desastre!», usamos el lenguaje para expresar emociones o actitudes, para mostrar nuestros sentimientos o estados de ánimo.

El Antirrealismo no cognitivista sostiene que concebir a los enunciados morales como pretendiendo describir un estado de cosas moral es consecuencia de no haber comprendido que el discurso moral no tienen ninguna pretensión descriptiva, sino que en cambio es usado, al igual que en los ejemplos apenas mencionados, para expresar actitudes. Desde este punto de vista, quien dice algo como: «Es inmoral interrumpir voluntariamente la vida del feto», como quien dice «Es una obligación moral que el Estado cobre impuestos a los que más tienen para mejorar la situación de los pobres», no hace más que expresar una actitud acerca de las acciones en cuestión. En el primer caso se expresa, por ejemplo, una actitud desfavorable hacia la interrupción voluntaria del embarazo, en el segundo caso se expresa una actitud favorable respecto de la acción del Estado indicada.

Dentro del Antirrealismo no cognitivista existe también una diversidad de variantes. Desde la más extrema de Alfred Ayer en *Lenguaje, verdad y lógica* (Ayer, [1936]1971), según la cual los enunciados morales expresan emociones.<sup>4</sup> Así, la única diferencia entre un enunciado como «Lo que usted hizo es un aborto» y un enunciado moral como «Lo que usted hizo al abortar está mal» es el tono. Como si el segundo enunciado fuese escrito en rojo, o con signos de exclamación o pronunciado con voz más severa. En la literatura filosófica se suele ilustrar esto mediante las siguientes fórmulas:

«Interrupción voluntaria del embarazo» Buuu!

«Cobro de impuestos por parte del Estado a los más ricos» Hurra!<sup>5</sup>

Otras variantes se alejan de este punto de vista. Por ejemplo, según la posición más sofisticada de Richard Hare (Hare, [1952]1975), en *El lenguaje de la moral*, a través de los enunciados morales se expresan preferencias (prescripciones) universalizables; es decir, aplicables a todas las personas alcanzadas por el juicio moral, incluida quien lo formula. Cuando digo «Está mal que abusos de la asistencia económica del Estado», no estoy simplemente alzando el tono de voz, sino que estoy expresando una preferencia universal que se traduce en la prescripción: «Ninguna persona (incluido yo mismo) debería abusar de la asistencia económica del Estado».<sup>6</sup>

Sobre esta base, el Antirrealismo no cognitivista puede ser definido del siguiente modo:

### (3) *Antirrealismo no cognitivista*

(i) Tesis semántica: los enunciados morales no son ni verdaderos ni falsos, sino que son usados para expresar actitudes (tesis no cognitivista).

(ii) Tesis ontológica: no existen hechos normativos.

---

4 En este breve texto no he abordado la relación entre las dos tesis usadas para caracterizar las posiciones morales analizadas. Aquí simplemente señalo que Ayer defiende la tesis semántica como consecuencia de su defensa de la tesis ontológica. Es decir, en el capítulo primero de *Lenguaje, verdad y lógica*, Ayer defiende el criterio de que un enunciado es genuinamente descriptivo si de su verdad o falsedad se sigue alguna diferencia observacional (tesis del empirismo radical). Dada esa tesis, Ayer procura en el capítulo 6 encontrar un lugar para los enunciados éticos (y los estéticos). Dado que en virtud de la tesis del empirismo radical, no pueden ser entendidos como describiendo algo (su verdad o falsedad no produce diferencias observacionales), entonces han de ser entendidos como simple expresión de emociones.

5 Con diferencias, la teoría emotivista de los enunciados morales también fue defendida por Charles Stevenson (1971).

6 Esta es una manera simple de distinguir entre el prescriptivismo de Hare y el emotivismo de Ayer, pero las diferencias son más profundas desde el punto de vista filosófico.



Para precisar la tesis no cognitivista, es necesario tener en cuenta la distinción entre enunciados que reportan estados mentales, emociones, preferencias o actitudes y enunciados que las expresan. Una cosa es decir que un sujeto *S* evalúa un objeto *O* de manera positiva o negativa y otra cosa es expresar (o mostrar) una evaluación.<sup>7</sup> Ejemplo del primer tipo: «(No) me gusta esto» y ejemplo del segundo tipo: «Esto es bueno(malo)».

El primer tipo de juicio de valor son aquellos enunciados en primera persona acerca de las actitudes propias de quien los formula. Su corrección, como la de cualquier enunciado descriptivo, depende de que coincida con el estado de cosas referido. En el caso de los enunciados acerca de actitudes, los criterios típicos son comportamentales, ya sea lingüísticos, tales como decir que algo es bueno o bello, ya sea no lingüísticos, como disfrutar o querer poseer el objeto, protegerlo contra el peligro o alejar otros sujetos que también lo desean (Von Wright, 2000:350).<sup>8</sup> En definitiva, se trata de enunciados contingentes que pueden ser verdaderos o falsos pero, estrictamente, no son enunciados morales. Esto es lo que diferencia al no cognitivismo del subjetivismo. Esta última es la posición que sostiene que los enunciados morales son descriptivos de estados mentales y, por lo tanto, tienen valor de verdad y algunos son verdaderos.<sup>9</sup>

El no cognitivismo sostiene, en cambio, que los enunciados morales no describen los estados mentales o actitudes de quien los pronuncia, sino que son usados para expresar (mostrar) las valoraciones del sujeto acerca de la cosa en cuestión. Así, enunciados como los siguientes: «Esto es bueno» o «Esto es aberrante» expresan las actitudes de aprobación y/o desaprobación de quienes los profieren y no dicen nada de que sea verdadero o falso. Se trata, según Von Wright, de un notable caso en el que la forma gramatical de un enunciado no refleja su forma lógica. Si bien se presentan en forma indicativa y poseen la forma gramatical típica de las afirmaciones, no afirman nada acerca del mundo.

Finalmente, Von Wright afirma que la teoría emotivista es, además, una buena interpretación de algunas afirmaciones de Wittgenstein. Así, en el *Tractatus* Wittgenstein sostiene que el discurso moral es un sinsentido, puesto que

---

7 Véase Von Wright (2000:351).

8 Ciertamente estos comportamientos pueden ser suprimidos (actitud sin comportamiento) o simuladas (comportamiento sin actitud). Pero, sostiene Von Wright, la supresión y la simulación presuponen las relaciones genuinas. Es decir, uno no puede simular o suprimir las propias emociones hacia ciertos objetos sin haberse antes familiarizado con los comportamientos asociados a la emoción en cuestión.

9 Por ejemplo, el enunciado «El Estado debe cobrar impuestos a los ricos para ayudar a los pobres» es verdadero si buena parte de la población de un país tiene una actitud favorable respecto de esa acción del Estado.

no se refiere a nada. De todas formas, sigue Wittgenstein, si bien no *dice* nada, *muestra* algo importante. Según Von Wright, para entender estas afirmaciones es necesario recordar que, para el Wittgenstein de *Tractatus*, *decir* algo equivale a formular un enunciado que puede ser verdadero o falso, es decir, un enunciado acerca de hechos. Dado que los hechos morales/normativos no existen, un enunciado como «Esto es bueno» no *dice* nada. No *afirma* nada.

Argumentos a favor del Antirrealismo no cognitivista.  
Concepción humeana de la acción y juicios morales

Uno de los argumentos más difundidos a favor de esta variante del antirrealismo es la facilidad con la que da cuenta de la capacidad motivacional de la moral. Por lo general consideramos que sostener sinceramente el enunciado moral «La tortura de personas inocentes es una aberración», implica estar motivado a actuar en consecuencia. Lo moral tiene, en términos de Stevenson, cierto magnetismo (Stevenson, 1937:16). No consideraríamos sincera a la persona que sostenga que ese enunciado es verdadero y, no obstante, torture inocentes. La moral motiva la acción. Pues bien, bajo una concepción estándar de la psicología humana, el no cognitivismo explica con sencillez este rasgo de la moral. Según la concepción estándar que suele ser asociada con las ideas de David Hume, existen dos clases principales de estados psicológicos: creencias y deseos. Las creencias pretenden describir el mundo y por ello son susceptibles de ser valoradas en términos de verdad y falsedad. Los deseos pretenden representar cómo debe ser el mundo, y por ello no son susceptibles de ser valorados en términos de verdad y falsedad. Esta imagen psicológica proporciona un modelo para comprender la acción humana. La acción humana es resultado de una combinación de ambos estados psicológicos que, en general, pueden ser denominados doxásticos (dirección de ajuste del estado mental al mundo) y connativos (dirección de ajuste del mundo al estado mental). Los estados doxásticos nos dicen cómo es el mundo y por lo tanto cómo ha de cambiarse para volverlo como lo piden nuestros estados connativos. De este modo, si afirmar que «La tortura de personas inocentes es aberrante» es un modo de expresar una emoción, preferencia o actitud contraria a la tortura ello explica que nos veamos motivados en ese sentido.

Por el contrario, dada la concepción estándar de la acción humana, el Realismo moral incluiría una visión inadecuada de la motivación moral. Ello en cuanto se vería obligado a sostener que la mera existencia de una creencia (moral) motiva, pero ello es incompatible con que la motivación depende de la existencia de un deseo.

## Un argumento en contra del Antirrealismo No Cognitvista. La comprensión de la acción valorativa

No hay duda de que algunas acciones son expresivas, es decir, tienen la finalidad de expresar actitudes. Ciertamente es posible realizar este tipo de acciones sin pronunciar palabras, es decir, es posible comunicar actitudes sin la mediación del lenguaje. Incluso, algunas veces, este tipo de comportamiento se acerca a acciones reflejas y es dudoso que se trate de comportamiento intencional. Pero otras veces, en cambio, resulta claro que algunas acciones que expresan actitudes son ciertamente intencionales.<sup>10</sup> Este es el caso de los actos lingüísticos realizados mediante la preferencia de enunciados morales. Formular un enunciado moral es un acto intencional. Es decir, no es algo que nos sucede, sino algo que hacemos. Por lo tanto, una teoría que procure explicar las acciones de los sujetos que profieren enunciados morales (mediante los cuales expresan su aprobación o desaprobación acerca de determinada acción) debe satisfacer las exigencias que, en general, se imponen a toda teoría explicativa de acciones. Este es precisamente el flanco hacia el que apunta la crítica de Terence Cuneo contra el antirrealismo no cognitvista. La tesis central de Cúneo es que «no es capaz de dar cuenta adecuadamente de las intenciones que, referidas al acto [lingüístico], poseen los agentes cuando se involucran en el discurso moral» (Cuneo, 2006:36).<sup>11</sup> En condiciones normales, un agente realiza el acto [lingüístico] A mediante la preferencia de un enunciado E si y solo si el agente tiene la intención de realizar A mediante la preferencia de E.<sup>12</sup> El argumento es que la preferencia sincera de enunciados morales se realiza con un tipo de intención incompatible con el tipo de acto que, según el no cognitvista, realizan los agentes involucrados en el discurso moral. Según la crítica, es falso que el acto llevado a cabo por el hablante al proferir un enunciado valorativo sea un acto de expresión de una actitud de aprobación o desaprobación hacia un estado de cosas u objeto no moral. Y es falso porque, según Cuneo, los agentes no tienen la intención de realizar

---

10 «Cierta tipo de comportamiento es *expresivo* —por ejemplo, el asombro, el disfrute, el disgusto, el llanto, el dolor—. (Se trata, en sentido amplio, de “actitudes emocionales”). Este comportamiento suele acercarse a reacciones reflejas. Por lo que es dudoso que sea intencional. Por ejemplo: me asusto y grito. Decimos “yo grito” y no “mi cuerpo produce el sonido del grito”. Pero si, como decimos, no pude evitar gritar (lo cual puede ser cierto), entonces es de dudar que el grito pueda serme atribuido. No soy “responsable” de él. Una razón por la que de todos modos es natural decir que fui yo quien lo hizo es, pienso, que gritar *puede* ser, de hecho normalmente es, claramente intencional» (Von Wright, 1998:101).

11 Cuneo se refiere a «actos ilocucionarios», pero no creo necesario introducir esta complicación en el presente trabajo.

12 Simplifico y ajusto a la terminología de este trabajo la formulación de Cuneo (2006:41).

ese tipo de acto, sino que tienen la intención de realizar un acto lingüístico descriptivo.

Para que el no cognitivismo sea verdadero respecto de la acción de quien profiere un enunciado valorativo, debería también ser verdadero que, en condiciones normales, cuando el agente profiere sinceramente un enunciado moral, no tiene la intención de describir, sino que tiene la intención de expresar una actitud.<sup>13</sup> Pero capturar lo que un hablante está intentando decir al involucrarse en un tipo de discurso exige tener en cuenta algunos hechos acerca del agente, entre los cuales se cuentan las convicciones del agente acerca de la naturaleza de ciertos rasgos de la realidad. Y, sostiene Cuneo, la evidencia muestra que los agentes al proferir un enunciado valorativo poseen la intención de afirmar o describir algo acerca de un estado de cosas u objeto moral. Casos como el del «creyente religioso tradicional» (un judío, un católico, un musulmán, entre otros) son de este tipo.<sup>14</sup>

Frente a esta crítica las posibilidades de respuestas son variadas, pero concentrándonos en las que consisten en ajustar alguna de las tesis del Antirrealismo no cognitivista podríamos señalar las dos siguientes: o bien sostener que los enunciados morales son efectivamente descriptivos, pero que de ello no se sigue, como sostiene el Realismo, que algunos sean verdaderos. O bien sostener que los enunciados morales son efectivamente una expresión de actitudes, pero que de ello no se sigue que el discurso en términos de verdad o falsedad no tenga sentido. Estas dos opciones dan lugar a las dos formas de Antirrealismo moral que quedan por explorar, a saber, el Antirrealismo del error y el Cuasirrealismo.

### **Antirrealismo del error**

Esta variante del antirrealismo se caracteriza por compartir la visión que del discurso moral tiene el Realismo, a saber, que se trata de una empresa con la pretensión de describir hechos morales y, por lo tanto, comparte una parte de la tesis semántica: que los enunciados morales poseen valor de verdad. Sin embargo, dado que rechaza la tesis ontológica, es decir, dado que sostiene que no hay hechos morales, concluye que lamentablemente la empresa moral es

---

13 De nuevo, simplifico y ajusto al presente trabajo la formulación de Cúneo: «En condiciones normales óptimas, cuando un agente realiza el acto oracional de proferir sinceramente un enunciado moral, el agente no tiene la intención de afirmar una proposición moral, sino que tiene la intención de expresar una actitud hacia un estado de cosas u objeto no moral» (Cuneo, 2006:43).

14 Estas afirmaciones se apoyan en evidencia sociológica relativa al discurso de los creyentes religiosos tradicionales (Cuneo, 2006:62–64).

la historia de un fracaso pues todos los enunciados morales, incluso si tienen la pretensión de describir, son irremediabilmente falsos. Y son falsos puesto que simplemente no existe el tipo de cosa que los podría hacer verdaderos. En realidad, el argumento es todavía más fuerte: porque es imposible que exista el tipo de cosa que los podría hacer verdaderos. En *Ética. La invención del bien y del mal* ([1977]2000), John Mackie introduce la versión más famosa de esta variante del Antirrealismo.<sup>15</sup> Para entender una posición de este tipo podemos recurrir a un dispositivo analítico propuesto por Bertrand Russell. Por ejemplo, el enunciado «El rey de Francia es calvo», dado que el Rey de Francia no existe y, por lo tanto, no refiere a nada en el mundo, ¿es un enunciado falso o un enunciado que no posee valor de verdad? Según Russell, la estructura lógica de ese enunciado revela que se trata de una conjunción de dos enunciados: «Existe un x tal que x es rey de Francia y x es calvo». Así, dado que el rey de Francia no existe, se trata de un enunciado falso, puesto que uno de sus elementos es falso.<sup>16</sup> Del mismo modo, quien afirma que un enunciado moral como «El Estado debe cobrar impuestos a los ricos para ayudar a los pobres», afirma dos cosas: (i) Existe un x, tal que x es un hecho moral y (ii) x impone al Estado el deber de cobrar impuestos a los ricos para ayudar a los pobres. Dado que no hay ningún x que sea un hecho moral, el enunciado es falso, ya que uno de los elementos de la conjunción lo es. La diferencia entre el enunciado sobre el rey de Francia y el enunciado sobre los deberes del Estado es que en el primer caso se trata de un error local, es decir, hay enunciados acerca de reyes que son verdaderos. En cambio, en el segundo caso se trata de un error global, ya que para esta variante del Antirrealismo el error se extiende a (casi) todos los enunciados morales (véase Mackie, [1977]2000:39 y 54). Y hay otra diferencia, central, y es que la falsedad de los enunciados sobre la realeza francesa es contingente (el mundo podría haber sido de tal forma que Francia todavía tenga una monarquía), mientras que la falsedad de los enunciados morales es necesaria, ya que los hechos morales son necesariamente inexistentes debido a las extrañas propiedades que deberían tener.

De este modo, las tesis que caracterizan a esta variante del antirrealismo son las siguientes:

---

15 Otras defensas de esta posición pueden verse en Joyce (2001); Olson (2014).

16 Véase Russell (1905).

(4) *Antirrealismo del error*

- (i) Tesis semántica: los enunciados morales tienen valor de verdad, pero son todos falsos (es decir, no hay enunciados morales verdaderos).
- (ii) Tesis ontológica: no existen hechos normativos.

Argumentos a favor del Antirrealismo del error.  
Mobiliario del mundo y hechos morales

Si bien, como adelanté, el Antirrealismo del Error acuerda con el Realismo moral acerca del discurso moral, ya que ambos son cognitivistas, es decir, para ambos los enunciados morales tienen la pretensión de describir hechos morales, sí desacuerdan acerca de la posibilidad de éxito de esta empresa. El error del Realista es creer que es posible que algunos enunciados morales pueden ser verdaderos al describir correctamente hechos morales. El error, más precisamente, es creer que tales hechos morales existen cuando, y este es uno de los argumentos centrales del Antirrealismo del error, es imposible que hechos como los necesarios para hacer verdadero un enunciado moral existan. No hay hechos morales en el mobiliario del universo.<sup>17</sup> Ello se debe a que hechos de ese tipo deberían poseer propiedades tales como una capacidad motivacional intrínseca completamente incompatibles con lo que sabemos acerca del mundo en el que vivimos, a partir de las adquisiciones de las ciencias naturales. Por ejemplo, en cierto momento de la historia de la ciencia, los científicos creían que la combustión se producía en virtud de la presencia de un elemento químico particular que poseen los objetos combustibles, a saber, el flogisto. En ese contexto, se produjo una buena cantidad de enunciados con pretensión descriptiva acerca del flogisto. Una vez que Lavoisier demostró que tal elemento no existe y que la combustión era fruto de una reacción química diferente, no quedó más que confirmar que todos esos enunciados eran falsos. El discurso sobre el flogisto es un discurso descriptivo pero falso. Algo análogo sucede con el discurso sobre hechos morales, con una diferencia: sabemos ya, dadas las adquisiciones de las ciencias naturales, que es imposible que tales hechos existan. A diferencia del Antirrealismo no cognitivista, el Antirrealismo del error está conceptualmente comprometido con la tesis ontológica. En el sentido que, si fuera a probarse que los hechos morales existen, esta versión el Antirrealismo debería aceptar que los enunciados morales pueden, a veces, ser verdaderos. En cambio, el no cognitivista tendría todavía

---

17 Este es el que Mackie denomina *argument from queerness* o «argumento de la singularidad» en la traducción española (Mackie, [1977]2000:42–46).

espacio para defender la tesis semántica, puesto que ella depende también de lo que hacen los usuarios del lenguaje.

El Antirrealismo del error puede sortear la crítica de Cúneo, pues logra acomodar las intenciones descriptivistas de quienes formulan enunciados morales y solo añade que esos enunciados son irremediabilmente falsos. Se podría pensar que resulta extraño que una teoría del error dé cuenta de las intenciones descriptivistas, pues esas intenciones incluyen también la pretensión de decir algo verdadero, al menos en algunas ocasiones. Es decir, parecería que puede explicar las intenciones mientras que quienes formulan enunciados morales no estén al tanto del error, pues una vez alertados sobre su error deberían, *ceteris paribus*, cambiar su comportamiento, tal como sucedió con el discurso acerca del flogisto. En otras palabras, una vez que los hablantes adviertan que, dado que no hay hechos morales, todos los enunciados morales son falsos, sería sospechoso que siguieran usando el mismo lenguaje en lugar de modificarlo para evitar el error. Nótese que no se trata de exigir al filósofo de la moral que muestre que los enunciados morales *son* verdaderos, sino que, al menos, muestre que es posible que lo sean. Sin embargo, incluso esta última parece una exigencia exagerada, pues llevaría, por ejemplo, a exigir al antropólogo que demuestre que el discurso acerca de la danza de la lluvia puede ser verdadero para afirmar que ha comprendido la práctica. Además, el partidario de la teoría del error podría intentar demostrar que hay algunas consideraciones, externas a la forma en que los hablantes conciben el funcionamiento de los enunciados morales, tales como la utilidad social, la eficacia en la persuasión, etc., que justificaría el uso continuado del discurso moral, incluso si los hablantes descubrieran o fueran ya conscientes del error.

## **El Cuasirrealismo**

El Cuasirrealismo puede también ser entendido como un modo de superar la crítica acerca de la comprensión de la acción valorativa. Para conseguir este objetivo propone afinar la relación entre la tesis semántica y la tesis ontológica acerca del discurso moral. Esto implica notar que, para usar con sentido enunciados morales, no es necesario que exista la entidad a la cual supuestamente el enunciado se refiere, como tampoco es necesario que tal entidad exista para que sea posible predicar verdad o falsedad del enunciado (Blackburn 1984:198). Para ello basta advertir dos características de los enunciados morales:

- (a) Son reflejos proposicionales de actitudes valorativas de quienes los pronuncian (Blackburn, 1998:77).

(b) Afirmar que un enunciado moral es verdadero, es simplemente lo mismo que afirmar que existe el deber impuesto por ese enunciado. A su vez, afirmar que existe el deber no es otra cosa más que expresar la propia actitud hacia la realización de la acción (Blackburn, 1998:77–83).

En pocas palabras, la primera característica implica que nuestras actitudes surgen como respuesta a propiedades del mundo que nos rodea (cfr. Blackburn, 1998:4). Así, lo que consideramos sensibilidad moral no es otra cosa más que una función de estímulos/creencias a respuesta/actitudes. Por ejemplo, cuando una persona A ve a otra persona B robando y exclama «¡Eso está mal!», el proyectivismo afirma que dado un conjunto de hechos, es decir, B llevando a cabo ciertas acciones, y la percepción por parte de A de esos hechos a partir de la cual forma ciertas creencias (acerca de lo que está haciendo B, de sus intenciones, de los efectos de lo que hace), se produce en A una emoción o actitud: de desaprobación frente a lo que hace B. Luego A proyecta esta emoción sobre el mundo y a ello sigue su afirmación de que «¡Eso está mal!». Esta es la razón por la que A «considera que el mundo contiene cierta propiedad, y el hacerlo no es consecuencia de rastrear una propiedad real del mundo, sino que es, en cambio, una reacción frente a una experiencia emocional» (Joyce, 2007).<sup>18</sup> Así, el Cuasirrealismo también sostiene que involucrarse en el discurso moral consiste en usar el lenguaje para expresar actitudes, previamente proyectadas en el mundo.<sup>19</sup>

¿En qué sentido puede usarse el término verdadero respecto de enunciados que expresan una actitud? Bueno, la verdad del enunciado de A es un producto de su propio trabajo intelectual y de su propia sensibilidad. Parafraseando a Blackburn, una persona *tiene que ganarse* su derecho a hablar de verdad moral. Dado que, en este contexto, la verdad tiene fuentes subjetivas, tales como actitudes, necesidades y deseos, su derecho depende de su capacidad para formar un conjunto mejorado de actitudes, donde la coherencia y la consistencia sean respetadas (Blackburn, 1984:197–202).

De acuerdo con Blackburn, afirmar que un enunciado moral es verdadero implica afirmar que la actitud reflejada pertenece al conjunto de actitudes que han superado todas las pruebas de mejoría, o al menos que no se consideran susceptibles de revisión. En palabras de Blackburn:

---

18 Véase Blackburn (1998:4–8).

19 «We project an attitude or habit or other commitment which is not descriptive onto the world, when we speak and think as though there were a property of things which our sayings describe which we can reason about, know about, be wrong about, and so on» (Blackburn, 1984:170–171).



La sugerencia simplemente es que definimos «el mejor conjunto posible de actitudes», como el conjunto de actitudes que han superado todas las oportunidades de mejora. Decir que un juicio evaluativo es verdadero es decir que es miembro de ese conjunto, o que se deriva de él. Si llamamos al conjunto  $M^*$ . Entonces, si  $m$  es un juicio evaluativo concreto, que expresa una actitud  $m$ :  $m$  es verdadero =  $m$  es miembro de  $M^*$ . (Blackburn, 1984:198)

Sobre esta base, el Cuasirrealismo respecto del discurso moral puede ser definido como sigue:

(5) *Cuasirrealismo moral*

- (i) Tesis semántica: los enunciados morales son usados para expresar actitudes, pero el discurso en términos de verdad puede ser conservado.
- (ii) Tesis ontológica: no existen hechos morales normativos.

Dado que, de acuerdo con el Cuasirrealismo, el discurso moral en términos de verdad y falsedad puede ser mantenido, este punto de vista logra también acomodar la gramática proposicional del discurso moral y, de esa forma, responder a la crítica sobre la comprensión de la acción valorativa.<sup>20</sup>

## **Un argumento en contra al Antirrealismo moral. Antiarquimedeanismo**

Las tres formas analizadas de Antirrealismo se apoyan en una distinción fundamental y bastante extendida dentro de la filosofía moral, a saber, la distinción entre dos niveles de discurso: ética y metaética. Se trata de una estricta división entre un discurso de primer orden, moral y comprometido, en el que se formulan enunciados morales y un discurso de segundo orden, meramente filosófico y no comprometido. Es en el discurso de segundo orden donde se llevan a cabo las actividades filosóficas típicas: el análisis conceptual y la determinación del estatus lógico de los enunciados de primer orden. Las dos tesis del Antirrealismo se ubican en este discurso de segundo orden, externo al discurso moral. Así, por ejemplo, la tesis no cognitivista según la cual los enunciados morales no describen nada, sino que expresan emociones, no es (no podría ser) a su vez, un enunciado moral, sino que es un enunciado conceptual o filosófico de segundo orden acerca de enunciados morales. Según

---

<sup>20</sup> Aquí, por razones de brevedad, me he detenido únicamente en la propuesta cuasirrealista de Blackburn. Sin embargo, otros autores defienden tesis similares, véase, por ejemplo, Gibbard (1990).

esta tajante división, por ejemplo, el enunciado (a): «La interrupción voluntaria del embarazo es un acto moralmente aberrante», es de primer orden y comprometido, perteneciente al dominio de la moral; y el enunciado (b): «El enunciado (a) no posee valor de verdad» o «El enunciado (a) es inevitablemente falso», es de segundo orden, conceptual y descriptivo, perteneciente al dominio de la filosofía.

Ronald Dworkin critica a quienes sostienen esta división entre enunciados comprometidos y enunciados meramente conceptuales respecto del discurso moral. Los llama arquimedeanos, como si creyeran haber encontrado un punto de apoyo por fuera de la moral para diluirla. La crítica de Dworkin al arquimedeanismo se basa en varios argumentos, aquí solo me detendré en el que apunta contra el estatus lógico de los enunciados de segundo orden. La conclusión de Dworkin es que no es posible defender una posición antirrealista acerca del discurso moral sin en algún punto presuponer la verdad de algún enunciado moral, lo que acorrala al antirrealismo en una contradicción o, incluso peor, en la inestable posición de quien serrucha su propia base. Ello en cuanto, claramente, las tesis del antirrealismo no pueden ser, a su vez, enunciados morales, de otro modo se vería defendiendo tesis que, o bien son inevitablemente falsas o bien carecen de valor de verdad. Por lo tanto, el antirrealismo está comprometido con que sus tesis no pertenezcan al dominio de la moral. Obviamente, como señalé arriba, esto es lo que el propio antirrealismo sostiene cuando distingue dos niveles de discurso. El punto de Dworkin es que las tres versiones del Antirrealismo fracasan en sostener esa distinción. Y fracasan porque, según este autor, sus tesis pueden ser siempre traducidas a un enunciado moral, o, al menos, a un enunciado acerca de la existencia de razones morales a favor o en contra de una acción.<sup>21</sup> Para mostrar este argumento, Dworkin propone analizar la siguiente lista de enunciados:

- (a) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es verdadero, porque las personas tienen derecho a disponer de su propio cuerpo.
- (b) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es falso, porque el feto es una persona desde el momento de la concepción.
- (c) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es falso, porque no existe un Dios y solo un Dios podría imponer deberes morales.

---

21 El argumento aparece ya en Dworkin (1986), pero está desarrollado con mayor detalle en Dworkin (2014:60–93).

Los primeros dos enunciados son claramente morales y expresan posiciones opuestas acerca del estatus de verdad del enunciado, también moral, «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas». El enunciado (c) también es un enunciado moral, pero escéptico, ya que niega que los enunciados morales puedan ser verdaderos en virtud de que no se dan las condiciones (existencia de Dios) para que puedan serlo. Ahora bien, sigue el argumento de Dworkin, consideremos el siguiente enunciado:

(d) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es inevitablemente falso, porque no existen esos extraños hechos normativos o morales que lo harían verdadero.

Este último sería, a diferencia de los tres anteriores, un enunciado filosófico no comprometido acerca de la ontología moral, que formularía un defensor del Antirrealismo del error. Sin embargo, sostiene Dworkin, se trata de un enunciado fácilmente traducible a un enunciado del mismo dominio que los anteriores. Así, según nuestro autor, (d) puede ser entendido como

(d<sub>1</sub>) No es obligatorio, ni está prohibido, ni está permitido que las mujeres interrumpan voluntariamente el embarazo hasta las primeras 12 semanas.<sup>22</sup>

Pero entonces, de (d<sub>1</sub>) se sigue

(d<sub>2</sub>) No existe una razón a favor o en contra de que las mujeres interrumpan voluntariamente el embarazo hasta las primeras 12 semanas.

Y (d<sub>2</sub>) es un enunciado del dominio de la moral, puesto que se refiere a la (in)existencia de razones a favor o en contra de una determinada acción.

Por su parte, el argumento contra el Antirrealismo no cognitivista se apoya también en un argumento de traducibilidad, pero en este caso respecto de las tesis que esta variante del antirrealismo pretende negar. Dado que Dworkin ataca aquí a toda posición que defienda la tesis según la cual «Los enunciados morales no poseen valor de verdad ya que expresan actitudes», alcanza también al Cuasirrealismo. La estrategia de Dworkin es mostrar que la tesis que el Antirrealismo se propone negar, a saber, la tesis cognitivista (i.) «Los enunciados morales son verdaderos o falsos», puede ser traducida a una tesis de primer orden, o perteneciente al dominio de la moral, por lo que su negación, argumenta Dworkin, pertenece también a ese dominio.

---

<sup>22</sup>Dworkin ofrece otra posible traducción: (4a) No es obligatorio ni está prohibido que las mujeres interrumpan voluntariamente el embarazo hasta las primeras 12 semanas. De lo que se seguiría, argumenta Dworkin, que la interrupción está permitida y esto último, claramente, es un enunciado moral. Sin embargo, esta traducción recibió críticas en un simposio previo a la publicación del libro y por eso propone también la traducción que indico en el texto. El simposio se encuentra publicado en la *Boston University Law Review*, 90(2), 2010.

Por ejemplo, la tesis cognitivista puede ser expresada y traducida de los siguientes modos:

(i.a.) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es objetivo.

Lo que se traduce en el enunciado de primer orden:

(i.b.) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» será verdadero incluso si nadie lo considerara verdadero.

O también en

(i.c.) La verdad del enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es universal.

Lo que se traduce en el enunciado de primer orden:

(i.d.) El enunciado en «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es verdadero para todos, sin importar en qué cultura o lugar se encuentran.

O también:

(i.e.) La verdad del enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» es absoluta.

Lo que se traduce en el enunciado de primer orden:

(i.f.) El enunciado «Existe el deber de permitir a las mujeres interrumpir voluntariamente el embarazo hasta las primeras 12 semanas» no es superado por otro enunciado moral.

Por lo que negar (i) implica negar (i.b.), (i.e.) y (i.f.), todos enunciados morales.

## Referencias bibliográficas

- Ayer, A.J. ([1936]1971). *Lenguaje, verdad y lógica* (trad. de Marcial Suárez). Martínez Roca.
- Blackburn, S. (1984). *Spreading the Word: Groundings in the Philosophy of Language*, Clarendon Press.
- Blackburn, S. (1998). *Ruling Passions*. Clarendon Press.
- Cuneo, T. (2006). Saying what we Mean. En Shafer-Landau, R. (Ed.), *Oxford Studies in Metaethics* (pp. 35–71). Vol. 1. Clarendon Press.
- Dworkin, R. (1986). *Law's Empire*. Belknap Press of Harvard University Press.
- Dworkin, R. (2014). *Justicia para erizos*. Fondo de Cultura Económica.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.
- Hare, R.M. ([1952]1975). *El lenguaje de la moral*. Universidad Nacional Autónoma de México.
- Harman, G. y Thomson, J.J. (1996). *Moral Relativism and Moral Objectivity*. Blackwell.
- Joyce, R. (2001). *The Myth of Morality*. Cambridge University Press.
- Joyce, R. (2007). Moral Anti-realism. *Stanford Encyclopedia of Philosophy*.
- Mackie, J.L. ([1977]2000). *Ética. La invención del bien y del mal*. Gedisa.
- Olson, J. (2014). *Moral Error Theory. History, Critique, Defence*. Oxford UP.
- Russell, B. (1905). On denoting. *Mind*, 14, 479–493.
- Stevenson, Ch.L. (1937). The Emotive Meaning of Ethical Terms. *Mind*, 46(181), 14–31.
- Stevenson, Ch.L. (1971). *Ética y lenguaje* (trad. de Eduardo Rabossi). Paidós.
- Von Wright, G.H. (2000). Valuations. *Ratio Juris*, 13(4), 347–357.

## 5. Escepticismo moral

*Leandro Paolicchi*

Hace ya algunos años, un pensador importante del siglo xx señalaba una particularidad que caracterizaba al estado actual de la ética (Apel, 1973). Por un lado, los avances científico-tecnológicos más importantes (por ejemplo, el desarrollo de armas nucleares) hacen necesarias un conjunto de normas que pongan en condiciones a los hombres de hacerse responsables por los riesgos que esos desarrollos producen sobre la población, no ya de una manera localizada (por ejemplo, en un país específico), sino ahora de forma *global*. Pero también, por otro, un tipo de pensamiento que está detrás de esos adelantos pone en entredicho cualquier posibilidad de fundamentar una ética, que por correspondencia no puede ser tampoco local sino que, por los *alcances* de los avances tecnológicos, debe ser también *global*. Es decir, la situación paradójica consistiría en que la ética es hoy más que nunca *necesaria* pero a la vez sus posibilidades de fundamentación y aplicación muy *difícultosas*.

Otro autor, cercano al anterior en términos de pensamiento ético, ha señalado (Habermas, 1983) que la moral no necesita de una justificación *definitiva* para funcionar y, por lo tanto, no tendría uno de los problemas señalados en el párrafo anterior. La confirmación de un punto de vista ético no es necesaria para que esa ética pueda funcionar en determinados aspectos del mundo contemporáneo. Ser reacios con respecto a la validez *última* de determinados principios no impediría que la sociedad pudiera funcionar con esos principios u otros. La gente no va por ahí preguntándose todo el tiempo por la fundamentación de los preceptos que respeta, los sigue (o no) y ya.

De todas maneras, reconoce también este filósofo, la moral encuentra innumerables obstáculos que no son los de una fundamentación pero sí los de incontables formas de escepticismo *práctico* que resisten formar parte de cualquier regulación. La ética encontraría, desde esta perspectiva, no sus mayores dificultades en el momento de su validación, sino en el momento de su implementación *práctica*. Los esfuerzos deberían estar, parece decirnos este autor, no tanto en demostrar la validez teórica de determinadas normas cuanto en asegurarnos que las personas las respeten mediante determinados mecanismos que no pueden sino incluir la *coerción institucional* en diversas formas.

Para otras corrientes de solo pensamiento, no solo la justificación de un conjunto de principios morales no sería necesaria sino que incluso la idea de un ámbito dedicado a cuestiones morales o de ética es un síntoma de que algo no anda bien dentro de una sociedad. Esta idea, que aparece bosquejada,

por ejemplo, en la primera generación de la Escuela de Frankfurt o en Marx (Stahl, 2017:505), parecería poner en juego una forma de escepticismo político o de sospecha *ideológica* sobre la totalidad de este ámbito de reflexión. Los problemas de la ética se resuelven de otra manera y, si hay una disciplina que habla de problemas éticos, es síntoma de un mal funcionamiento en la base de esa comunidad. Aquí el escepticismo ya no está dirigido contra diferentes empresas dentro de la ética sino al conjunto de un ejercicio disciplinar dedicado a esas cuestiones.

Lo que puede entreverse en los párrafos anteriores son diferentes formas de escepticismo que rodean y trasvasan a una disciplina como la ética y su objeto temático: la moral. Un escepticismo con respecto a la posibilidad de *fundamentar* normas morales, uno con respecto al rendimiento *práctico* de esas normas y otro con respecto a la *disciplina* como un todo. No obstante, como se verá más adelante, ellos no son los únicos y existen de la más variada gama dejando en evidencia que el escepticismo con respecto a diferentes aspectos de la ética (incluso a ella como un todo) y a las normas morales, están más extendidos que nunca en el mundo contemporáneo. Abordarlo en este artículo nos asegura siempre un abanico nutrido y actual de posiciones y temas.

Las condiciones socioculturales del capitalismo parecen ser las más propicias para la expansión de los más diversos estilos de suspicacia en general y moral en particular. La expansión de una racionalidad instrumental, asociada con la lógica del mercado capitalista, parecería ser la enemiga de cualquier intento de proponer un conjunto de reglas que pueden no ser funcionales a esta lógica. O mejor dicho, la idea de que las reglas del mercado capitalista son las únicas viables para establecer la convivencia entre los seres humanos es la culpable de que cualquier intento de limitarlas o reemplazarlas por otra caiga bajo sospecha. Como se mencionó al comienzo, la expansión de esa misma lógica en el ámbito científico intelectual es la responsable de que sea en extremo difícil dar fundamento a una ética transnacional absolutamente necesaria.

Si se realiza una genealogía un poco más amplia y se ubican los análisis en la transición de las sociedades tradicionales a las modernas pueden encontrarse allí dinámicas de *racionalización* que ponen en jaque un conjunto de tradiciones y cánones de comportamiento, disolviendo códigos éticos históricos e imponiendo al mismo tiempo estándares muy altos para las regulaciones que buscan establecerse en el mundo a partir de estas condiciones socioculturales. A la vez, esas dinámicas de racionalización van acompañadas por patrones de *individualización* que agregan obstáculos a la idea de códigos normativos que vinculen sin cuestión a grandes masas poblacionales. Puede decirse que, al calor de estas dinámicas históricas, la duda moral en términos normativos y también prácticos se ha garantizado un ámbito de supervivencia por mu-

chos años. Los esfuerzos hechos desde la ética normativa, y no solo desde ella, deben luchar contra una marea histórica verdaderamente grande y poderosa que arrastra olas desde muchos siglos atrás.

De todas maneras, el escepticismo no comienza ni con el capitalismo contemporáneo ni con la modernidad sino que tiene una larga tradición que es posible remontar hasta la antigüedad (Maliandi, 2004:106). No es objetivo de este trabajo hacer una historia *exhaustiva* y detallada de esta escuela sino más bien presentar algunos de los argumentos más fuertes que circulan en el panorama *contemporáneo*, así como de las formas más comunes de este pensamiento en el ámbito de la ética normativa y de la metaética actual. El resultado debería ser un mapeo básico de los argumentos más consistentes de escepticismo moral en las discusiones de los últimos años. Las condiciones históricas y filosóficas que limitan las posibilidades de la ética normativa y su aplicación, demarcan a su vez también las formas del escepticismo.

## Una cartografía inicial

Si, como se mencionó anteriormente, el escepticismo tiene modos muy diferentes, es necesario hacer operar sobre ellos alguna forma de reducción para poder entenderlos, discutirlos y sopesar su consistencia. Después de todo, no todas las figuras escépticas son sostenibles hoy en día y no vale la pena discutir las en el panorama contemporáneo, por lo que algún tipo de filtro que las ordene es una forma correcta de empezar. Algunos procedimientos de clasificación (que también son muy amplios) buscaron entenderlas como respuestas *específicas* a determinados problemas en el ámbito de la praxis humana. En este sentido, una de las clasificaciones posibles de las variedades de escepticismo (Maliandi, 1998:297) ha tendido a reducir las a *dos* modos básicos. Así se habló, por un lado, de un escepticismo de la *validez* y, por otro, de un escepticismo de la *vigencia*. Esta primera clasificación deja afuera por lo pronto alguna de las fórmulas escépticas sugeridas al comienzo; sin embargo, es útil para presentar las representaciones más usuales de la tradición que nos ocupa. Entre las primeras variedades (las de la validez) deben contarse aquellas posturas que sostienen que no puede *fundamentarse* de ninguna manera un conjunto de principios éticos y entre las últimas (las de la vigencia), aquellas que sostienen que los principios éticos solo valen para determinados grupos en *determinados* momentos pero no son universales ni de una manera sincrónica y menos diacrónica.

En el primer caso, la validez está entendida como la posibilidad de *justificar racionalmente* determinado conjunto de normas, independientemente de



cuántos las sigan o cuál es el alcance que realmente tengan dichas normas. Si son adoptadas solo por un conjunto de personas o han sido compartidas por grandes culturas no suma en nada a la *validez* que solo descansa en la posibilidad de justificar *racionalmente* el por qué esas normas deben ser aceptadas. En torno a la validez se ha dividido tradicionalmente la discusión en dos niveles (Maliandi, 1998:299). Por un lado, la validez abordada *desde la ética normativa* y, por otro, la abordada desde la *metaética*. Desde la ética normativa la discusión dará lugar a sus propias formas de escepticismo y desde la metaética la discusión también tendrá sus versiones escépticas particulares. Si bien la división entre ética normativa y metaética es siempre disputada, se mantendrá en este trabajo a los fines de simplificar la exposición.

Aunque existen múltiples variantes de escepticismo en metaética (Sinnott-Armstrong, 2006:9 y ss.), estas se dan *principalmente* alrededor de las posiciones que niegan cualquier validez a los *enunciados* de la moral. Desde este punto de vista, negar validez a las manifestaciones morales significa intentar demostrar que dichas expresiones no reflejan o describen ningún tipo de hecho moral objetivo y por lo tanto carecen de cualquier contenido *cognitivo*. Esta posición da lugar a las posiciones denominadas *no cognitivistas*. Contraponiéndose a ellas, están las que buscan rescatar algún tipo de validez para los enunciados de la moral y reconstruir a partir de allí algún tipo de contenido epistémico para sus enunciados, dando lugar a las posturas llamadas *cognitivistas*. Se volverá sobre este punto más adelante.

Otra forma de presentar el escepticismo en metaética ha consistido en dividirlo en dos formas principales, relacionadas con las mencionadas en el párrafo anterior. Por un lado, se encuentran aquellos que niegan *tajantemente* que exista algún tipo de *conocimiento moral*, entendido este como la posibilidad de que las creencias morales sean verdaderas y estén justificadas. A esto se lo denomina escepticismo *académico*. Por otro lado, se encuentran aquellos escépticos incluso con respecto a lo que afirman los escépticos académicos, es decir, la posibilidad de establecer *radicalmente* que no existe conocimiento moral alguno. Este último grupo duda inclusive que se pueda afirmar con *seguridad* la inexistencia de algún tipo de conocimiento moral, entendido como la posibilidad de que pueda haber alguna vez creencias morales verdaderas y justificadas. A este último grupo se los denomina «escépticos pirrónicos» (Sinnott-Armstrong, 2006:10-11).

La *vigencia*, por otro lado, se ha tendido a concebirla como el *seguimiento de hecho* de un conjunto de preceptos o reglas por parte de una cultura o población en algún momento de la historia. Naturalmente, cuando una sociedad respeta o actúa de acuerdo a una norma es porque de alguna manera cree que es *válida*. No obstante, es claro también que esto no siempre es el caso y

en las sociedades actuales muchas normas son seguidas independientemente de lo que se crea acerca de su validez (racional) y se la acata por cuestiones de algún tipo de beneficio que conlleva su seguimiento o porque existe un poder de sanción que castiga al que no siga esa norma.

Aquí resulta interesante ya desmarcar al escepticismo de su asociación con el relativismo con el que muchas veces se lo asocia. Pues si algunas formas de escepticismo niegan la posibilidad de fundamentación de normas morales o restringen su validez a determinados grupos y momentos, el relativismo implica una *fusión* de ambos momentos. El relativismo no niega ni la validez ni la vigencia (y por lo tanto no cabe llamarlo escepticismo), sino, como se ha sostenido (Maliandi, 1998:297), entiende la validez de determinados preceptos como aquellos que gozan de vigencia en determinados grupos o circunstancias. La validez se disuelve, por así decir, en la vigencia  *fáctica*. En el relativismo hay validez y vigencia, solo que de un modo acotado e indiferenciado.

## **El escepticismo en la ética normativa**

El escepticismo en este punto se presenta de una manera más radical que el relativismo, pues mientras que este último funde la validez en la vigencia, aquel *niega* la posibilidad de establecer la validez  *absoluta* de cualquier norma moral, ni siquiera mediante la vigencia. Las formas más paradigmáticas de esta postura es posible encontrarlas, como en muchos de estos casos, en la cultura griega y en particular en los sofistas. Se cita a menudo a Gorgias (Maliandi, 2004:106; Kuhlmann, 1984:545) como el iniciador de esta corriente de pensamiento, pero es recién con sus representantes contemporáneos que el escepticismo en ética normativa adquiere sus rasgos más notorios.

Es Nietzsche, también citado en este punto como el primero en distinguir las dos formas de negación presentadas al comienzo, un destacado iniciador contemporáneo de esta postura. Con respecto al escepticismo de la validez se suele traer a colación su mención a los «negadores de la moralidad» (Nietzsche, 2016:74). Allí ha dicho que están, por un lado, los que niegan que los hombres actúen  *realmente* por motivos morales y, por otro, los que niegan que los juicios morales puedan sustentarse en  *verdades*. Mientras que el primer inciso se refiere a un escepticismo de la vigencia, el segundo sería atribuible a uno de la validez. Ahora bien, tal como está formulado parecería que el escepticismo de Nietzsche en ese punto es un escepticismo de la validez  *aplicado a los enunciados morales*, es decir, claramente  *metaético* por lo que no voy a extender aquí sobre este punto y lo dejaré para la sección III.

Tal como se alcanzó a entrever al comienzo, el escepticismo en términos de ética normativa puede adoptar por lo pronto tres formas esquemáticas (Kuhlmann, 1984:545):

- 1) La fundamentación de normas no es necesaria.
- 2) La fundamentación de normas podría ser necesaria pero no es posible.
- 3) La fundamentación de normas es necesaria y posible pero no tiene ningún tipo de incidencia en la práctica real de las personas.

Con respecto al primer punto, esa postura aparece representada por aquellos que sostienen (por ejemplo, Herman Lübbe en *Philosophie nach der Aufklärung* de 1980) que los esfuerzos por fundamentar normas es una preocupación puramente *especulativa*, exclusiva de ciertos círculos *intelectuales* alejados de la praxis cotidiana de las personas. Los proyectos de justificación en términos prácticos son completamente inútiles, pues la praxis diaria de los individuos tiene una capacidad de regular los desacuerdos, y cuando estos se producen, rara vez se recurre a teorías normativas de justificación dudosa para su resolución. En algún punto, las prácticas morales cotidianas deben permanecer intactas pues allí los sujetos tienen claro qué es justo o ante qué son responsables.

Con respecto al segundo punto, la negación de la posibilidad de fundamentación de normas morales tiene una larga historia y representa una de las formas de escepticismo más extendidas. En la actualidad, uno de sus representantes más destacados proviene de lo que se llamó «racionalismo crítico». Si bien no se la ha comprendido explícitamente como una forma de escepticismo *ético*, el desarrollo de las ideas de Karl Popper (1980) y de Hans Albert (1980) en torno a la falibilidad de la razón tiene consecuencias similares sobre este tema. Si bien los racionalistas críticos no niegan alguna forma de justificación, sí se manifiestan en contra de cualquier tipo de fundamentación «última», por lo que quienes abrigan la esperanza dentro de la ética normativa de una demostración de este tipo ven en los racionalistas críticos algo parecido a los escépticos en términos de fundamentación. Tanto este escepticismo como el mencionado en el tercer punto serán discutidos más adelante en las secciones V y VII respectivamente.

## El escepticismo en la metaética

Como se ha dicho, al escepticismo ha tendido a vérselo no solo como una postura de ética normativa sino también como una postura específica dentro de la *metaética* y en este sentido como una respuesta concreta al valor de los *enunciados* de la moral. Aquí también han existido clasificaciones que han propuesto reducirlo a algunas formas básicas. Tal vez la cuestión más importante dentro de los problemas de la metaética sea la de si los enunciados morales representan alguna forma de *conocimiento*, es decir, si los enunciados morales reflejan algún tipo de hecho *moral objetivo* fuera del lenguaje y si pueden hacerlo de una manera correcta o no, es decir, si son *verdaderos* o *falsos*.

Las posturas que responden afirmativamente a esa cuestión se las ha denominado *cognitivistas* y a los que niegan tal posibilidad *no cognitivistas*. Es dentro de estas últimas donde pueden encontrarse las configuraciones más sofisticadas y complejas de escepticismo ético en los debates contemporáneos. En verdad esta discusión, que es una de las más importantes dentro de este ámbito, se da a menudo de la mano de arduas discusiones epistemológicas que no tienen que ver *directamente* con lo que se asocia normalmente con cuestiones éticas y menos con el escepticismo ético. Aquí el escepticismo adopta el cariz de una postura epistemológica con consecuencias para la ética, pero no directamente ética.

Tal vez la primera postura radical en negar cualquier tipo de contenido epistémico a los enunciados morales haya provenido del neopositivismo. Para los miembros del Círculo de Viena, cualquier término o enunciado al que no le correspondiera un correlato empírico en el mundo objetivo carecía de sentido. Aplicando este criterio, primero a la metafísica, pero también a la moral, el resultado estaba claro: las proposiciones de la moral carecen de sentido y por supuesto carecen de cualquier tipo de contenido epistémico. Son propiamente *pseudoproposiciones*.

Como contrapartida, la postura que ha sostenido *tradicionalmente* la posibilidad, no solo de *definir* los términos éticos en espacios de referencias empíricas, sino que estos pueden llegar a encarnar una forma de conocimiento ha sido el *naturalismo* (Maliandi, 2004:128; Sinnott-Armstrong, 2006:135-152; Hare, 1999:71-90; Tugendhat, 1997:65-77). Tal como su nombre lo indica, el naturalismo sostiene que los términos éticos pueden definirse en función de determinadas propiedades *naturales* del mundo. Como representante de esta corriente se suele mentar al utilitarismo, pero esto no está exento de controversia, pues existen otras clasificaciones que ve en el utilitarismo una forma de no cognitivismo (aunque débil) y, por lo tanto, una forma de escepticismo metaético (Habermas, 1996). Una enumeración más o menos extensa de

modelos de cognitivismo debe incluir no solo a los naturalistas sino también a los «transnaturalistas». Ambos serían representantes de lo que se conoce como «definicionismo» por lo que también habría no «definicionistas», como por ejemplo, los intuicionistas. Aquí la división pasa por la posibilidad de «definir» precisamente los términos en los enunciados morales en función de ciertas propiedades objetivas habilitando de esta manera que puedan ser considerados *verdaderos* o *falsos*. No obstante, toda esta enumeración solo sirve para poder contraponerla con esas formas de escepticismo metaético que estamos presentando.

Como se dijo, la forma más primitiva de no cognitivismo estuvo en un inicio representada por el Círculo de Viena, especialmente por Carnap. El filósofo alemán dio su interpretación de cuál era la característica básica de los enunciados de la moral y para él ese tipo de juicios no son asimilables a ninguna forma de *descripción*, sino que deben ser considerados en verdad «imperativos disfrazados». Cuando alguien dice «no debes robar» (por mencionar un ejemplo a menudo usado) lo que en verdad está queriendo decir es «no robes», es decir, está *ordenando* no robar. A esta posición primigenia se la conoce como «imperativismo».

Sin embargo, a esta primera forma de no cognitivismo siguieron otras posturas que fueron afinando su interpretación de estas expresiones para dar finalmente con una más adecuada que la de Carnap. Así, si bien insistían en negar cualquier tipo de contenido cognitivo a esos enunciados, trataban de ser fieles a lo que ellos creían que esas proposiciones representaban. De esta manera, como un desarrollo ulterior de la interpretación de Carnap surge la lectura que realiza Alfred Ayer, dando lugar a lo que se conoce como *emotivismo*. Para Ayer, los términos con los que hablamos de normas y deberes en realidad no son más que la expresión de sentimientos que experimentamos frente a nuestras acciones y la de los otros. De cualquier manera, el emotivismo adquirió una forma más sofisticada en la formulación de Charles L. Stevenson y en su distinción entre un significado descriptivo y uno emotivo.<sup>1</sup>

Un lugar destacado dentro de la tradición del no cognitivismo merece la interpretación realizada por R. Hare del lenguaje de la moral. Hare, crítico del emotivismo, cree que nuestro lenguaje moral en realidad es un reservorio de *prescripciones* y nada tiene que ver con la expresión de emociones ni con la manifestación de imperativos como creía Carnap. Por último, en este repaso veloz por las posiciones escépticas *clásicas* dentro de la metaética puede mencionarse a aquella que ha señalado que no es necesario remitir las expresiones

---

1 Para una exposición más amplia del emotivismo de Stevenson y su relación con otras corrientes de la ética analítica puede verse el artículo de Sábada citado en este trabajo.

de la moral a un solo tipo de función, sino que es posible rastrear muchas de estas funciones teniendo lugar al mismo tiempo en nuestro lenguaje. Esta posición, representada por Nowell-Smith y Warnock (Maliandi, 2004:134) sigue sin otorgarle ningún contenido epistémico a las expresiones morales, sin embargo, ha concedido que ellas pueden cumplir un conjunto de otros roles muy importantes en nuestro desenvolvimiento en el mundo de la vida cotidiana como las contempladas por las teorías anteriores.

### **El giro pragmático como nuevo terreno**

El giro pragmático (Böhler, 1986), concretado como una variante del giro lingüístico de principios del siglo xx, significó tanto un nuevo impulso para los intentos de fundamentación desde la ética normativa como para las tentativas de renovar alguna forma de cognitivismo metaético y reorganizó las discusiones de un modo original. Desde esta nueva perspectiva, se abrió la posibilidad de fundamentar principios éticos así como también de salvar algún contenido cognoscitivo para las expresiones morales en una discusión directa con algunas de las formas de escepticismo. La dimensión pragmática del habla cambió la manera de concebir la *objetividad* en el ámbito de las ciencias y por lo tanto alteró también la manera de interpretar el alcance de la *justificación* de las proposiciones morales. Si bien la dimensión pragmática del habla ya había sido considerada por alguno de los escépticos (Sábada, 1989:181 y ss.) no lo fue con la suficiente profundidad y sistematicidad como para trastocar la forma de concebir esos conceptos centrales.

Con este cambio en la objetividad y en la justificación se alteró también la idea de que para que pudiera concebirse a las manifestaciones morales como conocimiento debían necesariamente buscarse *hechos morales* que les correspondieran. La idea de objetividad entendida como *intersubjetividad* y la justificación entendida como *desempeño discursivo* (Habermas, 1996:53) abrió la posibilidad de rescatar un contenido epistémico en las expresiones morales a la vez que habilitó la reconstrucción de principios normativos que fungieran como principios morales.

El punto clave para este nuevo escenario es el reemplazo de una concepción semántica por una concepción *pragmática* de la fundamentación. La concepción semántica es, como se dijo, la que parecía dar marco a todas las formas de escepticismo vistas anteriormente y la que parece alimentar todavía hoy algunas posturas realistas (Sinnot-Armstrong, 2006:16 y ss.). Esta nueva noción pragmática consistiría en reemplazar la idea de que una proposición se considera fundamentada cuando es posible encontrar un correlato empírico que le

corresponda o cuando es posible derivarla deductivamente de proposiciones más básicas por otra en la cual una proposición está fundamentada cuando obtiene un acuerdo movilizado por razones (Habermas, 2004:312 y ss.).

Desde este nuevo punto de vista, la noción de un enunciado verdadero adquiere una significación inédita. Ya no remite inmediatamente a la confrontación con un material accesible solo en clave de percepción monológica sino ya a un juego de argumentación en donde de lo que se trata es de aportar y criticar *razones* que se movilizan en torno a la pretensión de validez de un enunciado. Validez significaría aquí, no suficiente material empírico aportado en favor de una proposición sino una suficiente cantidad de razones expuestas en torno a una pretensión de validez y un acuerdo finalmente sobre esa pretensión.

Para el caso de la ética o de los enunciados de la moral podría decirse que su validez (*corrección normativa* también se la ha llamado) no reside en la cantidad de material empírico que es posible movilizar para probar su contenido proposicional sino en la calidad de las razones y los argumentos que puedan esgrimirse para demostrar su respaldo. Supongamos que alguien dice «las mujeres deben recibir el mismo salario que los hombres a igual trabajo e igual responsabilidad». Desde un punto de vista tradicional, esta afirmación normativa no puede considerarse como descriptiva y por lo tanto no puede calificársela de verdadera o falsa. Es obvio que no hay y no puede haber un hecho moral que se corresponda con dicha afirmación que incluye un «debe» en su formulación. Esta afirmación visiblemente podrá considerarse válida o fundamentada en la medida en que los argumentos movilizados a su favor sean los mejores argumentos y puedan ser aceptados por el conjunto de las personas afectadas directa e indirectamente por las consecuencias de su implementación. Naturalmente, hay una gran cantidad de información empírica que puede aportarse para demostrar su viabilidad y su racionalidad y, por su puesto, esa validez no puede establecerse independientemente de esa cantidad de información, por ejemplo, referida al tipo de trabajo que realizan las mujeres y al tipo de pago que reciben en él, entre mucha otra información que existe ya en abundancia sobre este punto. De cualquier manera, lo que determina su aceptabilidad no es ese material empírico sino el *proceso deliberativo*.

Lo que puede observarse con esta nueva forma de concebir la justificación es que si bien las afirmaciones de la moral y las afirmaciones de tipo asertóricas son diferentes, en este nuevo escenario tienen procesos similares para determinar su aprobación. Se ha hablado incluso (Habermas, 1996:53-55) de un concepto *universal* de validez que puede aplicarse tanto al ámbito del mundo objetivo como al ámbito práctico social que es el que aquí nos interesa. Puede luego considerarse que el significado y las consecuencias de es-

tablecer la ratificación en uno y otro ámbito son diferentes en la medida en que los mundos objetivo y social son *ontológicamente* diferentes y por ello los enunciados tienen un sentido y un efecto completamente diverso en uno y otro ámbito. No obstante, la posibilidad de un procedimiento común de determinar el apoyo a ambos tipos de manifestaciones está abierta con el nuevo escenario que venimos proponiendo. El objetivo de cualquier manera no es desarrollar *in extenso* las posibilidades del cognitivismo sino rastrear las formas del escepticismo contemporáneo y sus mejores argumentos. En este sentido, y con todos estos cambios en el ámbito epistemológico, el escepticismo no podía ser el mismo y se modificó de una manera acorde. A nuevo ámbito, nuevo escepticismo.

### **El escepticismo luego del giro pragmático lingüístico**

El giro pragmático lingüístico pudo significar, como se dijo, la posibilidad de fundamentar principios en el ámbito de la ética normativa y por lo tanto un renovado desafío al escepticismo de la validez. Reconstruir y fundamentar principios éticos implícitos en el uso del lenguaje renovó la idea de una ética que puede ser universal en la medida en que haya seres humanos con la capacidad de hacer uso de sus habilidades lingüísticas. Naturalmente, el escepticismo debió elaborar nuevos argumentos, pues aquellos pensados tradicionalmente para los intentos anteriores de fundamentación ya no alcanzaban a este nuevo intento sobre la base de la performatividad del habla.

El escepticismo frente a esta nueva manera de fundamentación ética última y universal también fue diverso. Abarcó desde la imposibilidad de fundamentar de manera *última* principios en general, y morales en particular, a través de los presupuestos del habla, pasando por dudar del carácter *ético* de aquellos principios que se reconstruía en el lenguaje hasta el carácter *obligatorio* de estos (Paolicchi, 2014:141-208). A esto puede agregarse cierta sospecha en torno a que estas nuevas formas de fundamentación no hacen más que explicitar las intuiciones de una parte de la población mundial, más específicamente la europea, que ha transitado un desarrollo histórico posible (la modernidad) pero no el único y mucho menos el más legítimo. Esta es claramente una impugnación que esgrimen quienes ven en estos nuevos desarrollos una (nueva) forma de eurocentrismo (Allen, 2016).

Asimismo, también debe considerarse como una forma de escepticismo (presente también en el escenario pre giro pragmático) aquel tipo de comportamiento que, frente a la realidad de una moralidad que expresa algo más que meros intereses particulares, imperativos encubiertos, expresiones de senti-



mientos o prescripciones (o un poco de todo esto junto) y frente a la realidad de una fundamentación normativa de ciertos patrones éticos, a la vez que a la posibilidad de que estos patrones representen algún tipo de contenido epistémico, se refugia finalmente en un tipo de *acción* que escapa a cualquier tipo de regulación normativa. Esta última modalidad de escepticismo será abordada hacia el final del trabajo.

El escepticismo contemporáneo en materia de ética normativa se ha nutrido especialmente de una manera de objetar la fundamentación presentada en los escritos de Karl Popper y que Hans Albert trató de trasladar hacia la ética. Si bien sus pensamientos no se presentan a sí mismos como escépticos, han sido muy fructíferos para el desarrollo de estas nuevas formas de impugnación. La propuesta de Popper, continuada por Albert, ha sido caracterizada como racionalismo crítico porque a diferencia del llamado por ellos racionalismo «no crítico» o «comprensivo» (Popper, 1980, II:431) es consciente de sus propios límites y, por lo tanto, se muestra como una teoría *falible*. Desde este punto de vista cualquier fundamentación de principios no puede ser sino también precaria, es decir, no *última* y por lo tanto revisable todo el tiempo con ayuda del conocimiento de las ciencias. Como es sabido, y por ello no voy a extenderme demasiado sobre esta forma de escepticismo, estos autores han desarrollado la idea de que cualquier intento de fundamentación última recae necesariamente en lo que denominaron *Trilema de Münchenhausen*, es decir, en tres formas no viables de fundamentación. Una primera opción del trilema es intentar la fundamentación de una proposición por deducción de otra que se cree última, cayendo en un regreso *ad infinitum*. La segunda es la interrupción arbitraria de esta secuencia de deducción y, finalmente la tercera, es incurrir en una *petitio principii* al tomar como fundamentado aquello que se debía probar. Un punto interesante que se desprende de toda la argumentación de estos autores es el referido a la fundamentación del propio principio del racionalismo crítico. Pues bien, según la perspectiva de Popper no puede haber fundamentación última alguna y por lo tanto la opción por el propio principio y en última instancia la elección de la racionalidad debe ser una decisión irracional. Para Popper, la decisión entre la razón y su contrapartida irracional solo puede ser una decisión irracional, en su caso a favor de la razón.

La respuesta por parte de quienes propusieron la nueva manera de fundamentación consistió en decir que este escepticismo permanecía apresado en una idea semántica de justificación e ignoraba las posibilidades latentes en la consideración de los actos de habla y en la reconstrucción de los presupuestos que hacen posible *incluso* las mismas afirmaciones escépticas con respeto de la fundamentación, al contenido cognitivo de las afirmaciones morales o

finalmente *cualquier otra* forma de afirmación. La respuesta fue precisamente el desarrollo de la idea de fundamentación pragmática antes presentada. Es decir, si analizamos detenidamente las afirmaciones escépticas de Albert descubriremos una *estructura* que debe presuponerse de manera inexorable para que tales expresiones tengan sentido. La apuesta fuerte de las pragmáticas formales fue reconstruir y fundamentar esta estructura para *toda* argumentación con sentido (Kuhlmann, 1985:24). Esta armazón, compuesta por un conjunto de presupuestos irrefragables no puede negarse explícitamente pues se recae en una inconsistencia pragmática anulando el sentido de lo dicho. Es decir, no se puede rechazar en el plano proposicional lo que se está presuponiendo como *válido* en el plano pragmático. Así, como se ha encargado el mismo Apel de demostrar (Apel, 1998:48), la posibilidad misma de discutir el trilema de Münchhausen solo puede hacerse si se deja a un lado la dimensión pragmática del habla. Si se la tiene en cuenta, tanto Albert como Popper recaen en una inconsistencia pragmática, es decir, sosteniendo las afirmaciones del dilema contradicen aquello que dan presupuesto como *válido* en la dimensión pragmática de su formulación. Esta variante del escepticismo es insostenible.

### **El etnocentrismo como una forma de escepticismo de la vigencia**

Mientras que el escepticismo de la *validez* tiene sus dificultades serias cuando se trata de desarrollar su negativa a la posibilidad de fundamentar normas éticas luego del mencionado giro pragmático, el de la *vigencia* parece tener un punto más fuerte. Es más, puede decirse que con solo mirar prácticas culturales, políticas o empresarias, el escepticismo de la vigencia se mueve aquí sobre el terreno triunfante. Los que sostienen la validez universal de algunos preceptos éticos tienen que aferrarse a prácticas mínimas asentadas en instituciones con una incidencia muy puntual sobre determinados ámbitos. No obstante, cuanto más decisivas son las cuestiones y más poderosos los actores que intervienen, la ética debe ceder lugar a un fino juego de poder e intereses particulares.

Muchos de los que se han dedicado a la ética han temido este punto incluso más que el referido a la fundamentación. El panorama contemporáneo de la filosofía ha contribuido también, desalentando los esfuerzos por fundamentar en favor de una preocupación por el rastreo de cómo la moral se juega *a perder* en el tejido del poder y de las instituciones del mundo contemporáneo. Naturalmente, el escepticismo de la justificación ha sido también

un gran colaborador en este campo. Mientras las inconsistencias en torno a negar tanto la posibilidad como la necesidad de una fundamentación no han preocupado a muchos autores, sí los ha ocupado el rendimiento ético-político de los presupuestos del discurso. Podría decirse que una gran parte de las preocupaciones sobre este punto han girado en torno a un escepticismo de la vigencia que descrea de la posibilidad de trasladar la validez de los presupuestos igualitarios del discurso a las estructuras institucionales de las democracias contemporáneas.

A la vez, se ha ido perfilando junto con esta forma de escepticismo otra forma contemporánea más extendida que sostiene que cualquier intento de extraer principios de estructuras de la acción y del habla, y llevarlos a las instituciones de las democracias occidentales, esconde *en verdad* el intento de imponer determinadas intuiciones básicas de cierta parte del globo sobre las prácticas de otras culturas por considerarlas más importantes o centrales en el desarrollo de la humanidad. Concretamente, esta forma de escepticismo está muy atenta a la idea de transformar ciertas adquisiciones eurooccidentales en logros *globales* y con esa excusa oprimir prácticas culturales diferentes que no son funcionales al desarrollo del capitalismo. Esta variante del escepticismo diría: esas prácticas o patrones de comportamiento solo son válidas en el contexto europeo moderno. Las culturas que no han pasado por esas dinámicas no tienen que adoptar necesariamente esos preceptos normativos. La crítica al *eurocentrismo* de determinados proyectos normativos es el objetivo central de esta forma de escepticismo.

La manera que han tenido los recientes intentos normativos de defenderse frente a esta forma de impugnación es señalar que las estructuras que ellos reconstruyen y utilizan no son arbitrarias, sino que tienen tanta carga de *universalidad* como la que pueden reclamar para sí los procedimientos de la *ciencia*. Es decir, lo esgrimido por ellos es la utilización de los recientes descubrimientos de diferentes ciencias (la lingüística, la psicología, la antropología, entre otras) para poder a partir de allí obtener principios normativos.<sup>2</sup> Lo que aparece con esas prácticas científicas no son solo estructuras occidentales sino patrones *universales* del comportamiento y el habla del hombre *en general*. Que muchas de esas aportaciones se realicen en ámbitos europeos o anglosajones no debe desviar la atención de que son actualizaciones de lógicas *inherentes a lo humano* y no específicas de determinada cultura. Si bien las prácticas científicas son situadas, los resultados son *universales*.

---

<sup>2</sup> Lo que sobrevuela aquí es, por supuesto, el fantasma de alguna forma de falacia naturalista. Al respecto, puede verse Apel (1998:649–699).

La respuesta de quienes acusan de eurocentrismo a muchos de los desarrollos normativos europeos apunta a discutir la base de esas definiciones, es decir, la ciencia. El primer punto, tal vez evidente, en señalar por parte de quienes descreen de estos «logros» es que se trata de ciencias que tienen como objeto hombres concretos e históricos, es decir, individuos en sociedades *específicas*. No es la investigación del mundo *natural* que puede reclamar para sí un ámbito de legalidad más indiscutida que el de las ciencias sociales. Las investigaciones mencionadas sobre patrones culturales, desarrollos cognitivos y lingüísticos son producto de procedimientos propios de las ciencias sociales con los resultados falibles que ellas producen. Específicamente, muchas de las objeciones esgrimidas apuntan a mostrar que los resultados de esas investigaciones son altamente discutibles. Por mencionar un ejemplo: sin ser especialistas en las respectivas disciplinas, los resultados de las investigaciones de Piaget o Chomsky, que en su momento se pensaron como logros seguros, se encuentran hoy bajo profundas revisiones. Es decir, podemos pensar en algún momento que las teorías en torno a desarrollos evolutivos sobre patrones cognitivos son irreversibles y tiempo después darnos cuenta de lo discutible de sus resultados.

Las consecuencias de esto son bien claras, poner en la base de desarrollos normativos los logros de la ciencia no nos protege de incurrir en cierto eurocentrismo *residual*, sobre todo cuando esas ciencias se refieren al hombre y sus entornos socioculturales. Esto no significa que la ciencia no pueda ser utilizada para orientarnos en la búsqueda de patrones normativos, más bien es imprescindible en esta tarea. No obstante, no podemos utilizarla como refugio último para sostener una pretensión de universalidad *fuerte*.

### **Una línea de escepticismo práctico**

El escepticismo como forma de pensamiento es muy amplio y puede esgrimirse casi en cada rincón sobre el que la razón se extienda. Frente a todas las opciones que se barajaron anteriormente y que podrían hacernos dudar de su viabilidad en esos ámbitos, este tiene todavía lugares en los cuales puede refugiarse. Frente a las posibilidades que existen luego del giro pragmático de poder extraer un conjunto de normas de contenido ético y frente a la reconstrucción en esas normas procedimentales de una capacidad dadora de validez similar a la de los enunciados asertivos (y con ello cierto contenido epistémico), el escéptico podía todavía desconfiar de que esos procedimientos o normas solo fueran las intuiciones de culturas centrales del capitalismo, como la europea o la anglosajona, y su desarrollo respondiera al intento de imponer

esos patrones culturales sobre prácticas diversas y no funcionales al capitalismo. Los recursos de las ciencias pueden ser movilizados para mitigar esta sospecha, pero están siempre sometidos a las revisiones de una comunidad científica que también está prevenida de extrapolar rápidamente resultados locales sobre todas las sociedades. El peligro del eurocentrismo siempre acecha a los proyectos normativos procedentes del viejo continente.

De todas maneras, es posible considerar finalmente dos formas más de escepticismo. Una primera, con respecto al *valor* de las teorías éticas y una última con respecto a toda forma de normatividad por parte de los actores sociales que solo *actúan* rechazando cualquier regulación. La primera, podríamos decir, es todavía un escepticismo *intelectual* en la medida en que las dudas expresadas provienen de sectores que se dedican a cuestiones prácticas (Lariguet, 2015). La segunda, en cambio, podríamos denominarla un escepticismo *práctico* (Sinnot-Armstrong, 2006:11), pues no implica ninguna forma de argumentación o discusión ética sino más bien una *actitud* que rehúye toda normatividad, cualquiera sea la forma en que esta se establezca. Podría pensarse que no están relacionadas en absoluto, pero no es del todo así. En muchas ocasiones, un escepticismo del segundo tipo lleva a uno del primero.

Con respecto al alcance de las teorías éticas y de la normatividad en general, algo se ha mencionado al inicio de este trabajo. No debe subestimarse ese tipo de desconfianza pues es una de las más extendidas. Ello depende en gran medida, no solo de las propias teorías éticas, sino también ampliamente de la segunda forma de escepticismo mencionado en este apartado. El escepticismo con respecto al alcance de las teorías éticas está determinado no solo por la existencia de individuos que rehúyen cualquier forma de normatividad, sino por *patrones sociales* que resisten las regulaciones de tipo ético. Dinámicas estructurales del capitalismo que no tienen en cuenta preceptos ético-normativos de ningún tipo lleva a replantearse *radicalmente* el valor de la ética y de las teorías normativas que se proponen. Es decir, es el escepticismo propiamente *práctico* el que pone el foco sobre las teorías éticas, promueve una aprensión con respecto a la ética en general y pone en evidencia que la implementación de propuestas normativas depende en gran parte de condiciones sociales y culturales que no maneja.

La figura del escepticismo práctico ha sido trabajada también por Habermas, en un principio bajo la idea de un escéptico que no argumenta (1983:109 y ss.). Pero mientras que en esos primeros escritos sobre ética trataba de responder con estructuras del mundo de la vida ordenadas en torno a formas simples de comunicación de la cual el escéptico no podía escapar, en sus trabajos posteriores se vio obligado a considerar configuraciones institucionales como una forma de contener y mitigar lo que él denomina la *debilidad mo-*

*tivacional* de la moral. Es decir, en producciones posteriores Habermas se vio forzado a considerar un problema más *estructural* que se dejaba en evidencia con el escepticismo práctico y que afectaba también a la ética como teoría en cuanto tal. En *Facticidad y Validez*, Habermas reemplaza esa salida problemática de recurrir a las estructuras comunicativas del mundo de la vida por una complementariedad del *derecho*, pues ahora el «enemigo» ya deja de ser el escepticismo o lo es de una manera un poco velada.

El punto de partida para pensar en estructuras *institucionales* que aseguren el seguimiento de la moral es la *debilidad motivacional* que define a este tipo de normas. Las condiciones socioculturales que son propicias para la expansión de las diferentes formas de escepticismo son las mismas que cargan a la moral con una debilidad inherente que solo se hace viable en contextos institucionales asegurados por el derecho. Volviendo al comienzo de este trabajo, la transición moderna hacia sociedades capitalistas, la pérdida absoluta de peso de la tradición hace a la moralidad débil en cuanto a su *predisposición a la acción*. La moral, ya en un escenario pos convencional, es un saber más entre otros que «flota en el plano de la cultura» (Habermas, 1992:145-146). Sin una sanción institucional que la fije en normas del derecho, la moral contemporánea y con ella la relevancia de la ética, es una presa constante de alguna de las formas de escepticismo que se presentaron anteriormente.

## Referencias bibliográficas

- Albert, H. (1980). *Traktat über kritische Vernunft*. Mohr.
- Allen, A. (2016). *The End of Progress. Decolonizing the normative foundations of critical theory*. Columbia University Press.
- Apel, K.O. (1973). *Das Apriori der Kommunikationsgemeinschaft und die Grundlagen der Ethik*. En *Transformation der Philosophie* (pp. 358-435), II Bde. Suhrkamp.
- Apel, K.O. (1998). *Auseinandersetzungen in Erprobung des transzendentalpragmatischen Ansatzes*. Suhrkamp.
- Böhler, D. (1986). *Pragmatische Wende*. Suhrkamp.
- De Zan, J. (2002). *Panorama de la ética continental contemporánea*. Akal.
- Habermas, J. (1983). *Diskursethik – Notizen zu einem Begründungsprogramm*. En Habermas, J., *Moralbewusstsein und kommunikatives Handeln* (pp. 53-125). Suhrkamp.
- Habermas, J. (1992). *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Suhrkamp.
- Habermas, J. (1996). *Eine genealogische Betrachtung zum kognitiven Gehalt der Moral*. En Habermas, J., *Die Einbeziehung des Anderen. Studien zur politischen Theorie* (pp. 11-64). Suhrkamp.

- Habermas, J. (2004). Richtigkeit versus Wahrheit. Zum Sinn der Sollgeltung moralische Urteile und Normen. En *Wahrheit und Rechtfertigung. Philosophische Aufsätze* (pp. 299–346). Suhrkamp.
- Hare, R. (1999). Ordenando a la ética. Una clasificación de las teorías éticas. Ariel.
- Kuhlmann, W. (1984). Ist eine philosophische Letzbegründung moralischer Normen nötig? En Apel, K.O. et. al. (Eds.), *Funkkolleg: Praktische Philosophie/Ethik: Dialoge* (pp. 545–571), Bd. I. Fischer-Taschenbuch Verlag.
- Kuhlmann, W. (1985). Reflexive Letztbegründung, Untersuchungen zur Transzendentalpragmatik. Alber.
- Lariguet, G. (2015). Dilemas, escepticismo moral e imparcialidad. Consideraciones en torno a un prólogo de René González de la Vega sobre mi libro *Encrucijadas morales*. *Ápeiron. Estudios de Filosofía*, (2), 136–155.
- Lübbe, H. (1980). Philosophie nach der Aufklärung. Von der Notwendigkeit pragmatischer Vernunft. Econ Verlag.
- Maliandi, R. (1990). Anerkennung oder Argumentation? Zur Erörterung einer scheinbare Antinomie in der gegenwärtigen Ethik. En Apel, K.O. y Pozzo, R. (Eds.), *Zur Rekonstruktion der praktischen Philosophie. Gedenkschrift für Karl-Heinz Ilting* (pp. 283–303). Stuttgart-Bad Cannstatt: Frommann-Holzboog.
- Maliandi, R. (1998). La ética cuestionada. Prolegómenos para una ética convergente. Almagesto.
- Maliandi, R. (2004). Ética: conceptos y problemas. Biblos.
- Nietzsche, F. (2016). *Morgenröte. Gedanken über moralischer Vorurteile*. Holzinger.
- Paolicchi, L. (2014). Praxis, sentido y normatividad. Hacia una reconstrucción pragmática de las acciones humanas. Prometeo.
- Popper, K. (1980). *Die offene Gesellschaft und ihre Feinde*, II Bde. Francke Verlag.
- Sábada, J. (1989). Ética analítica. En Camps, V. (Ed.), *Historia de la ética* (pp. 162–220), T. III. Crítica.
- Sinnot-Armstrong, W. (2006). *Moral Skepticism*. Oxford University Press.
- Stahl, T. (2017). The Metaethics of Critical Theories. En Thompson, M. (Ed.), *The Palgrave Handbook of Critical Theory* (pp. 505–522). Palgrave Macmillan.
- Strawson, P. (2008). *Freedom and Resentment*. Routledge.
- Tugendhat, E. (1997). *Lecciones de ética*. Gedisa.

## 6. El relativismo metaético: argumentos a favor y en contra<sup>1</sup>

Lucas E. Misseri

El enfrentamiento tiene varios nombres: absolutistas contra relativistas, tradicionalistas contra posmodernos, realistas contra idealistas, objetivistas contra subjetivistas, racionalistas contra teóricos de las construcciones sociales, universalistas contra contextualistas, platónicos contra pragmáticos. *La verdad: guía de perplejos*, Simon Blackburn

Lo relativo es aquello que se opone a lo absoluto, por lo cual uno podría pensar que existen tantas formas de relativismo como de absolutismos. Por ejemplo, así como existe un relativismo epistemológico, que cuestiona el hecho de que exista una verdad absoluta, hay un relativismo ético,<sup>2</sup> que cuestiona la posibilidad de un principio absoluto en materia moral. Es importante comenzar definiendo al absolutismo ético porque, en muchos sentidos, el relativismo ético surge como reacción ante él y especialmente contra algunos de sus excesos:<sup>3</sup>

*Absolutismo ético*: postura según la cual existen verdades morales objetivas, de las que pueden derivarse principios aplicables a todos los seres racionales y en todo período histórico.<sup>4</sup>

---

1 Agradezco a Anna Bugajska y Santiago Truccone por sus valiosas sugerencias a una versión preliminar de este capítulo y a Guillermo Lariguet y a Rafael Buzón por la recomendación de bibliografía al respecto.

2 En la bibliografía pertinente se usa indistintamente el adjetivo «moral» o «ético», siendo incluso más frecuente que el término aparezca como «relativismo moral». Sin embargo, sigo aquí la distinción entre moral y ética que propone Maliandi (2004). Según esta, el término «moral» hace referencia al conjunto de creencias y prácticas en torno a lo bueno y lo correcto y el término «ética», a la reflexión filosófica en torno a la moral. Como el relativismo que me interesa encarna un cierto grado de complejidad filosófica prefiero el adjetivo «ético».

3 Quizá entre los excesos podrían incluirse formas acríicas de absolutismo que derivan en fanatismo, etnocentrismo e intolerancia contra otras formas de vida que son ajenas a esos principios percibidos como absolutos.

4 Para esta definición me baso en una combinación de las de Bunnin y Yu: «The view that there are certain objective moral principles which are eternally and universally true, no matter what consequences they bring about» (2004:451); y Rozuel: «Moral truths do exist, independently from our personal feelings or beliefs, and independently of social norms. These moral truths generate principles that are absolute and apply to all at all times» (2013:1052).



De la negación de los distintos aspectos de esta postura pueden derivarse varias alternativas asociadas, en parte, a algunas formas de relativismo. Por ejemplo, puede negarse que existan verdades morales y aquí es donde en metaética se habla de posturas no-cognoscitivistas, es decir, que niegan que lo que hacemos cuando enunciamos juicios morales sea emplear proposiciones, o sea que no se puede aumentar a través de ellas nuestro conocimiento con enunciados calificables como verdaderos o falsos. De esto se siguen, no solo posturas no cognoscitivistas como el emotivismo y el prescriptivismo, sino también la claudicación ante la posibilidad de todo conocimiento ético: el escepticismo.

En general, el relativismo no niega que pueda hablarse de verdades morales, sino más bien niega el carácter objetivo de dichas verdades. Esto es problemático porque la idea de verdad subjetiva parece poner en jaque la idea misma de verdad<sup>5</sup> y se entiende por qué, en muchos sentidos, el relativismo se ve asociado al escepticismo. No obstante, para salvar ese problema algunos relativistas como Gilbert Harman prefieren hablar de validez en lugar de verdad. De la imposibilidad de que haya un marco de referencia «absoluto» para la ética no se sigue que no pueda haber marcos relativos en los que haya principios válidos.<sup>6</sup> De esta visión se sigue una crítica al aspecto objetivista del absolutismo («verdades morales objetivas») y al aspecto universalista («principios aplicables a todos»). Si se opone por exceso al absolutismo, se opone también por defecto al nihilismo ético, es decir, a la creencia de que de la negación de la imposibilidad de un marco ético absoluto se siga la irrelevancia de la ética. Con esto ya se estaría en condiciones de dar una definición con la que trabajar en las siguientes secciones:

*Relativismo ético*: postura según la cual la validez de los principios morales depende de un cierto marco de referencia como el de una cultura, un grupo o incluso un individuo. Esa dependencia implica la negación de la posibilidad de un marco de referencia privilegiado.<sup>7</sup>

---

5 Para un abordaje de este problema, véase Blackburn (2006).

6 Algunos autores hablan también de contextualismo ético, para la distinción entre este y el relativismo, véase Moreso (2009).

7 Esta definición se inspira en las de Bunnin y Yu: «The view that ethical terms and principles are relative to cultures, societies, and even persons. There are different ethical judgments about the same subject, and there is no decisive method of reasoning that can adjudicate between these conflicting judgments» (2004:226). Rozuel: «There is no such thing as moral facts or moral truths; rather the boundaries of morality are determined by individuals or groups of individuals who agree on what is acceptable behavior and what is unacceptable. Moral judgments or moral values are always relative to a particular framework, for instance, one's culture, one's personal beliefs, or the context of the situation» (2013:1051).

De aquí se pueden distinguir varias formas del relativismo ético. El relativismo ético asociado al relativismo cultural, un relativismo ético asociado a ciertos grupos, y finalmente un relativismo ético asociado a cada individuo en lo que da en llamarse subjetivismo ético. En lo que sigue se evitará la referencia al relativismo cultural porque el término cultura es demasiado vago. Las culturas son muy difíciles de delimitar con precisión y aun cuando esa dificultad se pudiera salvar, estas incluyen otros tipos de valores además de los morales, como por ejemplo los valores estéticos, económicos, etc. Tampoco me ocuparé del subjetivismo ético. Mi foco estará puesto en el relativismo ético de grupos —con más de un individuo— y los desafíos que tiene defender ese tipo de postura.

El relativismo ético así definido puede ser dividido en tres tipos distintos, pero conciliables: el relativismo ético descriptivo, el relativismo ético normativo y el relativismo metaético. Si bien aquí se tratará principalmente del último, es importante dar cuenta de qué se entiende por los otros dos, dado que los mismos influyen en varios de los argumentos que se expondrán más adelante.

El *relativismo ético descriptivo* ha sido impulsado desde fines del siglo xix y principios del xx, principalmente a partir de los trabajos de antropólogos culturales estadounidenses como Willam G. Sumner y Franz Boas. Este relativismo consiste básicamente en la afirmación de la tesis empírica según la cual los distintos grupos humanos que pueblan el mundo sostienen, de hecho, creencias y prácticas morales distintas. Esta es una tesis que, si bien tiene sus matices, es ampliamente aceptada en la actualidad. Las diferencias muchas veces se notan especialmente en el contraste entre algunos países más modernizados, en los que se da prioridad al individuo sobre la comunidad, y los países más tradicionales que dan prioridad a valores comunitarios por encima de los individuales.<sup>8</sup> Esta forma de relativismo enfatiza la diversidad cultural humana y, en su forma más acentuada, se apoya en la creencia de que existen amplios desacuerdos en materia moral entre los seres humanos de esas diversas culturas. Una postura abiertamente opuesta a ella es el universalismo ético descriptivo, es decir, una posición que resalta los elementos compartidos en todas las culturas como por ejemplo cuando au-

---

<sup>8</sup> Véase por ejemplo el mapa cultural mundial de R. Inglehart y C. Welzel en el que se comparan dos ejes de acuerdo con distintos tipos de valores en los diferentes países del mundo de los que se tienen datos en la World Values Survey. En su versión de 2008, los autores contrastan en el eje x valores de supervivencia con valores de autoexpresión, y en el eje y valores tradicionales con valores seculares—racionales. Quedando así, por ejemplo, en el extremo de valores seculares de autoexpresión países de culturas cristiano—protestantes como Suecia y en el extremo de valores tradicionales de supervivencia países islámicos como Jordania. Si bien hay controversia sobre la exactitud del lugar de cada país en el mapa, hay bastante acuerdo sobre la diversidad de valores en juego en las distintas culturas del mundo.

tores como Judith Jarvis Thomson (Harman y Thomson, 1996) y Josep Joan Moreso (2009) brindan ejemplos de valores o juicios morales que parecerían gozar de aceptación universal.

Por su parte, el *relativismo ético normativo*<sup>9</sup> es muy poco usual, probablemente debido a las distintas dificultades argumentales que supone su defensa, que implican en general algún grado de autocontradicción performativa, es decir, contradicción entre la verdad de lo que se sostiene y la verdad asociada a lo que se hace al sostenerlo. A partir de la tesis empírica relativista ético descriptiva de la antropología cultural, los defensores de esta postura sostienen, no solo que existe diversidad de creencias y prácticas morales, sino una tesis normativa según la cual no debe interferirse, ni deben juzgarse las prácticas de otras comunidades con creencias morales distintas a las nuestras. Sin embargo, autores como Chris Gowans (2019) consideran que la tesis ética normativa no es una forma de relativismo *per se*, sino que se trata de una tesis que se añade al relativismo ético descriptivo o al metaético, dado que este no es independiente de aquellos. En cualquier caso, a esta forma de relativismo se oponen tanto el absolutismo como el universalismo ético, los cuales afirman que, aun cuando, de hecho, existan desacuerdos morales, esto no implica que no haya soluciones racionales objetivas para los mismos.

Finalmente, el *relativismo metaético* sostiene, en su forma más general, que la verdad o falsedad de los juicios morales es relativa, es decir, que depende del grupo que elabora dicho juicio. Asimismo, se establece que no hay tal cosa como un estándar objetivo que permita establecer que ciertos sistemas o conjuntos de creencias y prácticas morales sean superiores a otros. Esto supone varios problemas para otros presupuestos que también gozan de cierta aceptación global y que son cruciales para el mantenimiento de la tolerancia y el respeto por la diversidad humana, como es el caso de los derechos humanos. A esta forma de relativismo se oponen el universalismo y el objetivismo metaéticos, que en muchos aspectos forman parte de los presupuestos de discursos como los que van ligados a esos derechos humanos.<sup>10</sup> En vistas de la importancia de esta discusión para la moral y el derecho se abordarán a continuación algunos argumentos a favor del relativismo metaético subrayando algunas de sus debilidades.

---

9 Moser y Carson (2001) prefieren hablar de relativismo moral de los requisitos porque consideran que la tesis detrás de esa línea relativista implica la satisfacción de ciertos requisitos morales. No obstante, aquí se prefiere la expresión relativismo ético–normativo porque está implicada la misma tesis, pero el término es más acorde a las divisiones tradicionales de la ética en descriptiva, normativa y metaética.

10 Para una crítica al carácter absoluto de los derechos humanos y su reinterpretación como derechos *prima facie*, véase Laporta (1987).

## 1.

Los principales argumentos a favor del relativismo moral metaético se basan, por un lado, en la idea de diversidad de las culturas y, por el otro, en la tolerancia como un valor necesario y ligado al relativismo. Esto en parte se condice con lo que Emrys Westacott (2020) denomina el argumento negativo y el argumento positivo en favor del relativismo.

El argumento negativo expuesto brevemente sostiene que hay que ser relativista metaético porque ninguna teoría no relativista es satisfactoria. Este argumento subraya que hay desacuerdos en cuanto a cómo explicar la semiosis y el alcance de los enunciados morales que las otras teorías metaéticas, especialmente las absolutistas, objetivistas y universalistas, no han podido zanjar de modo concluyente. De la controversia, este argumento deriva la aserción de que efectivamente no hay un marco absoluto, porque de lo contrario podría haberse solucionado concluyentemente el problema metaético. Pero no necesariamente se trata de un buen argumento, porque por ejemplo del hecho de que existan partidarios del terraplanismo no se sigue que no pueda establecerse si la Tierra es o no plana, es decir, de la controversia no se sigue la no existencia de buena teoría o de ausencia de respuestas razonables a ciertos problemas. Sin embargo, este argumento hace hincapié en la brecha entre los enunciados descriptivos y los enunciados normativos. Los primeros pueden ser proposiciones y pueden ser verificables o falsables, mientras que los enunciados normativos que caracterizan al discurso moral y a las teorías ético-normativas no tienen esa posibilidad, sino que pueden ser correctos o incorrectos, válidos o inválidos. Pero la corrección o validez no se puede establecer de modo absoluto, sino que depende del marco de referencia del grupo en cuestión.

En cambio, el argumento positivo sostiene que el relativismo metaético, es decir, la negación de marcos de referencia absolutos en materia moral supone uno de los mejores argumentos a favor de la tolerancia y hasta una garantía para la misma. Del hecho de que sea imposible asumir un marco de referencia absoluto, se sigue la tesis ético-normativa de que uno no debe imponer el propio marco de referencia moral a los demás, porque no hay nada que lo avale para tal acción. Según este argumento esa falta de absolutez en materia ética quita fuerza a toda postura que busque imponerse a los demás como privilegiada y, por tanto, debe reconocer los otros marcos de referencia como igualmente válidos. Como se dijo más arriba, esta asociación entre relativismo y tolerancia puede estar originada en la experiencia histórica en la cual en muchos casos se evidenció cómo la intolerancia buscó su fundamento en alguna forma de absolutismo ético. Asimismo, hay algunos estudios empíricos

que correlacionan el relativismo con mayor tolerancia hacia las opiniones diferentes de la propia, cuestionando la intuición de que en general las personas tienden a ser absolutistas éticos.<sup>11</sup>

Un autor como el alemán Gustav Radbruch defendía esta tesis positiva en su *Der Relativismus in der Rechtsphilosophie* de 1934, cuando hacía derivar del relativismo la mayoría de los aspectos positivos del derecho moderno. Para Radbruch (2009:4–8) del relativismo se derivaban: en primer lugar, el positivismo jurídico, dado que al no haber una moral absoluta el derecho más que de verdad moral se trataba de autoridad, de voluntad; en segundo lugar, el liberalismo entendido como respeto por las libertades negativas del individuo; en tercer lugar, un derecho penal especial que permitiese penar los delitos por convicción; en cuarto lugar, el Estado de Derecho como Estado sujeto a su propia ley; en quinto lugar, la división de poderes como algo derivado del Estado de Derecho; en sexto lugar, el Estado democrático como un marco formal para la igualdad surgida de la ausencia de ese marco absoluto; en séptimo y octavo lugar respectivamente, la tolerancia general y un cierto grado de socialismo entendido como respeto por las libertades positivas. Es interesante, como se verá en la siguiente sección, que autores contemporáneos como Moreso (2012) se esfuerzan, desde una perspectiva rawlsiana, de mostrar exactamente lo opuesto a lo sostenido por Radbruch, esto es, que no hay una relación de necesidad entre relativismo ético, positivismo jurídico y liberalismo político. Una posible explicación de esa diferencia tiene que ver con cómo interpretar el relativismo radbruchiano. Según Rodríguez Paniagua (1963), hay que entender el relativismo del iusfilósofo alemán como una forma especial de relativismo que él denomina jurídico más que ético. Para este autor el relativismo sería un método más que una postura ética, en favor de esta interpretación puede citarse que el propio Radbruch considera que hay un «método relativista representado en la filosofía jurídica alemana por Max Weber y Georg Jellinek, Hans Kelsen y Hermann Kantorowicz» (2009:3). Sin embargo, se incluye aquí el argumento de Radbruch porque el mismo también puede ser defendido para la ética, aunque es cierto que la tesis subyacente a la fórmula que lleva su nombre es difícilmente compatible con el relativismo ético.

Entre los defensores más recientes del relativismo se encuentran los filósofos estadounidenses Gilbert Harman, Jesse Prinz y David Wong, quienes asocian al relativismo metaético con el naturalismo, es decir, con las ciencias naturales y su objetividad. En el caso de Harman, este parte de la compara-

---

11 Me refiero a Pölzler y Wright (2020). Sin embargo, debe hacerse notar que lo testeado por estos autores entra en el marco de la denominada «folk metaethics», es decir, las intuiciones metaéticas de personas que no se dedican al estudio de la ética ni tienen pretensiones de establecer una teoría acabada o coherente de la misma.

ción entre la física relativista de Einstein y la ética relativista. Harman cree que, así como la masa de un objeto depende de su marco espacio-temporal en física, del mismo modo lo correcto depende del marco de referencia que se asume. Por lo cual define y defiende al relativismo como la postura según la cual «no hay una única moral verdadera. Hay muchos marcos morales diferentes, ninguno de los cuales es más correcto que otro» (Harman y Thomson, 1996:5). Esta definición le permite diferenciar su postura de lo que considera extremos, por un lado, del absolutismo ético, según el cual solo existe una única moral verdadera, y por el otro, del nihilismo ético, según el cual de la negación de una única moral verdadera se sigue el desinterés por todo juicio moral. Ante el nihilista, Harman sostiene que la moral es importante como lo es el derecho, porque permite resolver conflictos. Pero frente al absolutismo y al derecho natural, Harman sostiene que no hay nada previo a la negociación (*bargaining*) que se establece entre las partes que forman los acuerdos (*conventions*) que conforman el derecho y la moralidad (1996:7 y 22). Harman construye su argumento a partir del relativismo ético-descriptivo, para luego hacer el salto al argumento que Westacott llama negativo y que, en vistas de que el absolutismo no puede dar cuenta de la diversidad moral adecuadamente, el relativismo es la mejor explicación (10) o la más plausible (63).

Es interesante que Harman se vale del prestigio de la física y de metáforas espaciales para explicar su relativismo, pero no obstante deja fuera elementos de la teoría física einsteiniana que no le resultan útiles: como por ejemplo el hecho de que la velocidad de la luz es una constante, es decir, es algo absoluto en la teoría de la relatividad. Pero, no obstante, eso no anula el argumento de Harman, sino que pone en duda el uso retórico de su comparación. Lo que es más importante es que describe a los distintos grupos de personas y sus creencias como si fuesen cosas tan diferentes como dos cuerpos celestes, como si no hubiese nada común en los seres humanos que pudiera establecer alguna clase de constante. En cualquier caso, el argumento negativo pone la carga de la prueba en el absolutismo. En palabras del propio Harman: «¿Por qué le parece (a algunas personas) que hay hechos objetivos no-relativos acerca de lo correcto y lo incorrecto?» (1996:13). Para el filósofo estadounidense no se trata de cómo deben ser las cosas o de lo que quiere decir una persona cuando hace un juicio moral, sino de cómo son las cosas en materia moral y cómo no son (17). Para él en general nuestro lenguaje tiende a ser cuasi-absolutista y lo que hacemos es proyectar nuestro marco de referencia moral a todos los demás como si fuera un marco absoluto, aun cuando no lo es. Intenta no hacer valoraciones sobre si hay que actuar de otra manera, sino más bien busca describir qué se hace y desde qué marco se llevan adelante los acuerdos que establecen la moral de un grupo.

Como se dijo más arriba, para Harman, las cosas en materia moral son relativas, pero eso no impide que haya algunos acuerdos morales desde los cuales ordenar nuestras relaciones en la sociedad. Esa moral por convención es reconocida como relativa a esas negociaciones entre los valores de las distintas partes involucradas y gracias a ella

Harman intenta eludir el problema de si deben ser tolerados o no aquellos que no la respeten. Su respuesta es tajante: «Los relativistas morales no necesitan ser tolerantes de los daños cometidos por criminales que no han aceptado la moral convencional» (1996:58). Pero el problema parece ser que en Harman la moral se confunde con la política, es solo el juego de fuerzas. Ahora los «criminales» podrían ser aquellos que quedaron fuera del juego de poder que establece la moral social o el derecho positivo. Por momentos parece que el discurso del relativismo hace de la moral un avatar de la política, del juego de fuerzas entre grupos de interés.

Algo similar ocurre con el caso de Richard Rorty (1996), quien partiendo del problema del relativismo epistémico y, en especial de las críticas a la concepción de la verdad del pragmatismo, considera que hay que reemplazar la idea de objetividad por la de solidaridad. Este autor cuestiona las versiones realistas en materia epistemológica para defender una interpretación del antirrealismo que él denomina antirrepresentacionalismo. Esa postura sostiene que no hay representaciones de la realidad sino construcción de hábitos. De modo que el relativismo ético no es entendible en el sentido de Harman, de la construcción de un marco de referencia para la moral, sino como ciertos hábitos de preferencia por el propio grupo. Esto que él denomina solidaridad y que sería compatible con el liberalismo, pero un obstáculo al etnocentrismo. Es decir, nuevamente otro autor intenta defender el vínculo entre una cierta forma de entender el relativismo como asociado al liberalismo. Asimismo, también aparece la idea de convertir a la ética en política, al derivar los problemas morales a la cuestión de la preferencia por los intereses de un cierto grupo sobre otros. A continuación, se verán algunas debilidades de este tipo de argumentos y el riesgo que pueden entrañar para con la tolerancia que dicen defender.

## 2.

Entre los argumentos en contra del relativismo ético se pueden distinguir aquellos que rechazan el argumento negativo, aquellos que rechazan el argumento positivo y aquellos que rechazan las consecuencias que se siguen —no solo prácticas, sino también teóricas— de sostener el relativismo metaético.

Aquellos que niegan el argumento negativo, se esfuerzan en mostrar que se exageran las diferencias entre los grupos humanos y que pueden encontrarse algunos candidatos a valores objetivos y universales. Quienes niegan este argumento no rechazan que haya diversidad de costumbres y valores, pero consideran que esa diversidad no es suficiente como para negar un cierto mínimo común básico compartido entre todos los seres humanos. Partidarios de este argumento son Gewirth y Rawls con su defensa de las necesidades básicas como un criterio objetivo y universal, aunque ellos se refirieron a bienes necesarios para un fin y bienes primarios respectivamente (Zimmerling, 1990:43). Asimismo, el ya citado iusfilósofo español (Moreso, 2009 y 2012) usa un ejemplo que más que atender a un conjunto de necesidades busca una proposición moral: «la tortura de niños por mera diversión es moralmente reprochable». Según Moreso, ese enunciado es absolutamente verdadero o al menos un buen candidato a serlo, dado que quien no lo aceptase como verdadero no sería una persona razonable en términos rawlsianos. Lo que replican relativistas como Harman es que ahora el problema está en definir lo razonable y como ellos niegan que haya un marco común para acordarlo en el contexto moral, el argumento sería insuficiente.

En cambio, autores como Atienza (2017) buscan desactivar la crítica subyacente al objetivismo deshaciéndose de aquellos aspectos absolutistas que el relativista condena, en especial la intolerancia producto del dogmatismo, e intentan mostrar que la tolerancia solo tiene sentido desde lo que con Fishkin denomina «objetivismo mínimo». Esta postura vendría a decir que hay algo valioso para todos los seres humanos más allá de sus perspectivas subjetivas, pero más acá del absolutismo ético. Ahora bien, Atienza parece dejar abierto el problema de la fundamentación metaética de ese objetivismo mínimo que parece solaparse con algunas formas de constructivismo en su defensa de lo objetivo como lo razonable. Antes de él, otros autores como Rawls (1980), Nino (1989) y Korsgaard (1996) se decantaron por alguna forma de constructivismo metaético,<sup>12</sup> es decir, pasaron el problema de la objetividad no a una realidad moral ajena a los seres humanos que estos pudieran captar, sino a la idea misma de intersubjetividad. Incluso algunos como Nino (1989:15) sostuvieron contra el argumento de Radbruch de la sección anterior, que es el constructivismo ético —y no el relativismo— el que conduce al liberalismo político que garantiza la tolerancia.

---

12 Para un completo recuento de las distintas formas de constructivismo en metaética y las dificultades que todas ellas enfrentan, véase Bagnoli (2020).



Por su parte, aquellos que niegan el argumento positivo lo hacen desde dos puntos de vista distintos. Por un lado, están los que consideran que si aquello que hay de positivo en el relativismo es el valor de la tolerancia, ese valor sería considerado como un valor universal por los relativistas y, por tanto, no sería relativo, sino consustancial a todos los grupos humanos. Esto hace a la postura relativista una posición contradictoria que se vuelve contra sí misma.<sup>13</sup> Pero, por otro lado, están también quienes van aún más lejos y sostienen que es necesario no ser relativista para que haya efectiva tolerancia como es el caso de Popper (1985), Moreso (2012), Arienza (2017) y la mayor parte de los pensadores partidarios del liberalismo igualitarista, yendo implícitamente contra la tesis de Radbruch antes mencionada. Muchas veces suele citarse la llamada «paradoja de la tolerancia» descrita por Karl Popper según la cual cabe preguntarse: ¿qué tiene que hacer el tolerante frente a aquel grupo con el que comparte un espacio vital que no incluye entre sus valores a la tolerancia? O en términos aun más directos, ¿hay que tolerar a los intolerantes? Parecería que si no toleramos a los intolerantes estamos siendo igualmente intolerantes, pero si los toleramos estamos mermando las bases de nuestra propia existencia y con ella la de nuestra jerarquía axiológica. Esto se ve claramente en política donde a algunos partidos extremistas se les impide la participación como representantes porque, de obtener el poder, lo que harían sería quitarles derechos a los otros grupos. Ahora bien, parecería así, que la tolerancia no puede apoyarse en el relativismo porque este atenta contra el mismo valor de la tolerancia. Si todo en materia de moral es relativo, la tolerancia misma es un valor relativo, de lo contrario parece ser un valor absoluto. Aun entre aquellos relativistas no académicos (*folk relativism*) parece estar marcado un fuerte presupuesto de que hay ciertas precondiciones que deben ser cumplidas, entre ellas la de aceptar las posturas ajenas. Pero ¿por qué? La respuesta a esa pregunta no parece más que ser una invitación a abandonar el relativismo, el reconocimiento de precondiciones básicas para la discusión de cualquier postura no requiere abrazar al absolutismo, pero sí alguna forma de objetivismo.

En última instancia están quienes sostienen que defender el relativismo implica aceptar que «todo vale». Esto porque si no hay valores que trasciendan lo relativo a cada grupo, no hay posibilidad de juicio crítico en torno a las prácticas disvaliosas de otras culturas, ni siquiera dentro del propio grupo. Un ejemplo límite que suele presentarse es el de la ablación de clítoris en algunos países (Kopelman, 1994). Aceptar el relativismo implicaría aceptar que no hay posibilidad de reprochar que esa práctica sea llevada a cabo a niñas en

---

13 Blackburn llama a este el «argumento rebotante» (2006:51) porque es un argumento que se vuelve en contra del relativista, «rebota».

contra de su voluntad porque es algo aceptado como moralmente encomiable en el marco de esa comunidad. Pero este relativismo dificulta la crítica, no solo entre diferentes culturas, sino entre diferentes períodos históricos de una misma cultura, por ejemplo, un estadounidense que se precie de relativista tendría serias dificultades para explicar por qué la esclavitud imperante en su país en el siglo XIX es reprochable sin apelar a algún valor trascendente o a algún grado de comparabilidad. Algo similar también podría ocurrir en el marco de una misma cultura, en algún sentido cuando se habla de una cultura se subrayan las semejanzas, pero se omiten muchas diferencias que también afectan a los valores. Se puede hablar de «cultura latina» para referirse tanto a los habitantes del sur del Río Bravo como hasta aquellos de la remota Ushuaia, pero puede haber fuertes diferencias entre una comunidad del norte y otra del sur incluso en una misma región, entre una urbana y otra rural, entre un grupo con altos ingresos económicos y otro con bajos ingresos, entre uno que sigue una religión y otro otra o ninguna. Sin embargo, pese a los desacuerdos en materia moral en nuestro día a día consideramos que tenemos bases para juzgar las conductas de aquellos que no piensan exactamente igual que nosotros, que podemos dar razones de la (in)corrección de esas conductas. Muchas de nuestras instituciones se apoyan en esa idea. Esto no quiere decir que no haya conflicto, no obstante, el camino hacia la solución, disolución o gestión de los conflictos parece dado más bien por la búsqueda de puntos en común como son por ejemplo las apelaciones a la razonabilidad, a los argumentos de razón pública o de bien común, independientemente de las cosmovisiones en las que nos encontramos. La búsqueda de un punto no relativo parece una necesidad de nuestras mismas prácticas morales.

### 3.

En este breve capítulo se ha dado una definición de relativismo como la postura filosófica que reacciona ante el absolutismo y que cuestiona al universalismo y al objetivismo éticos. Se lo ha clasificado en ético descriptivo, ético normativo y metaético. Se han ofrecido dos argumentos a favor del relativismo metaético (el negativo y el positivo) y tres en contra (la réplica al negativo, la réplica al positivo y uno combinado), procurando mostrar que incluso el valor de la tolerancia no se puede proteger desde una postura relativista.

En general, se considera que el relativismo cumplió una función interesante en la historia del pensamiento ético, mostrando brechas en los marcos absolutistas éticos y abriendo camino a una mayor sensibilidad a las particularidades de los distintos grupos. Por lo cual en un cierto aspecto una actitud

relativista en algún momento de la reflexión moral puede tener sentido y hasta ser saludable en tanto que fuerza a revisar las propias creencias morales. No obstante, el relativismo no puede ser la fuente de justificación de nuestras instituciones, ni una teoría ética en sentido fuerte que busque guiar nuestras acciones, porque se topa con las propias dificultades que ella misma pone al absolutismo. El relativismo impide buscar acuerdos entre las personas tolerantes al rechazar la posibilidad de que exista un marco común desde el cual establecer acuerdos, por lo cual deja el camino abierto a los intolerantes, a quienes no les interesa buscar acuerdos sino imponer sus condiciones a los demás.

En resumen, una actitud exploratoria de tipo relativista puede ser provechosa en una cierta etapa del razonamiento ético, hasta puede ser una virtud epistémica que invite a tener en cuenta los desacuerdos, las posturas ajenas a la propia y a revisar con ojo crítico las propias creencias. Pero nunca puede ser el punto de llegada del razonamiento ético porque, como tal, implica el abandono del propio razonamiento y, donde la razón cesa, se abren camino otros medios como la fuerza.

## Referencias bibliográficas

- Atienza, M. (2017). *Filosofía y transformación social*. Trotta.
- Bagnoli, C. (2020). Constructivism in Metaethics. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/constructivism-metaethics/>
- Blackburn, S. (2006). *La verdad: guía de perplejos*. Crítica.
- Bunnin, N. y Yu, J. (2004). *The Blackwell Dictionary of Western Philosophy*. Blackwell.
- Gowans, C. (2019). Moral Relativism. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2019/entries/moral-relativism/>
- Harman, G. y Thomson, J. (1996). *Moral Relativism and Moral Objectivity*. Blackwell.
- Inglehart, R. y Welzel, C. (2008). *Cultural World Map*. World Values Association. <http://www.worldvaluessurvey.org/WVSContents.jsp>
- Kopelman, L.M. (1994). *Female circumcision/genital mutilation and ethical relativism*. Second opinion, 20(2), 55–71.
- Korsgaard, C.M. (1996). *The Sources of Normativity*. Cambridge University Press.
- Laporta, F. (1987). *Sobre el concepto de derechos humanos*. Doxa: Cuadernos de Filosofía del Derecho, (4), 22–46.
- Maliandi, R. (2004). *Ética: conceptos y problemas*. Biblos.
- Moreso, J.J. (2009). *Una (relativa) refutación del relativismo moral (A (Relative) Refutation of the Moral Relativism)*. Análisi e Diritto, 2009, 8997.
- Moreso, J.J. (2012). *Positivismo jurídico, relativismo moral y liberalismo político*. Teoría Política, Annali, (2), 103–110.
- Moser, P.K. y Carson, T.L. (Eds.) (2001). *Moral Relativism: A Reader*. Oxford University Press.

- Nino, C. (1989). *El constructivismo ético*. Centro de Estudios Políticos y Constitucionales.
- Pözlner, T. y Wright, J.C. (2020). *Anti-Realist Pluralism: A New Approach to Folk Metaethics*. *Review of Philosophy and Psychology*, (11), 53–82. <https://doi.org/10.1007/s13164-019-00447-8>.
- Popper, K.R. (1985). *La sociedad abierta y sus enemigos*. Orbis.
- Radbruch, G. (2009). *Relativismo y Derecho* (trad. de Luis Villar Borda). Temis.
- Rawls, J. (1980). *Kantian Constructivism in Moral Theory: Rational and Full Autonomy*. *Journal of Philosophy*, 77(9), 515–572.
- Rodríguez Paniagua, J.M. (1963). El relativismo jurídico de Radbruch y su consecuencia política. *Revista de Estudios Políticos*, (128), 77–99.
- Rorty, R. (1996). *Objetividad, relativismo y verdad* (trad. de Jorge Vigil Rubio). Paidós.
- Rozuel, C. (2013). Ethical Absolutism Versus Ethical Relativism. En Idowu, S.O.; Capaldi, N.; Zu, L.; Gupta, A.D. (Eds), *Encyclopedia of Corporate Social Responsibility* (pp. 1051–1058). Springer.
- Westacott, E. (2020). Moral relativism. En *Internet Encyclopedia of Philosophy*. <https://www.iep.utm.edu/moral-re/>
- Zimmerling, R. (1990). *Necesidades básicas y relativismo moral*. *Doxa: Cuadernos de Filosofía del Derecho*, (7), 35–54.

## 7. Universalismo y particularismo

Graciela Vidiella

Una de las cuestiones que no ha dejado de interesar a la ética occidental desde sus inicios es si el razonamiento y el juicio moral se orientan por principios o si se enfocan exclusivamente en los contextos y casos particulares. Los filósofos universalistas piensan que se determinan —o deberían hacerlo— por principios universales válidos para todas las situaciones.

Puede considerarse que el primer universalista fue Platón con su teoría de las virtudes objetivas e invariables que, *qua* ideas, trascienden el mundo empírico. Su discípulo Aristóteles, en cambio, al proponer el término medio como definición de las virtudes éticas, hace de la percepción del contexto el objeto del juicio moral:

El que se desvía poco del bien no es censurado, tanto si se excede como si peca por defecto. Pero sí lo es el que se desvía mucho, porque éste no pasa inadvertido. Sin embargo, hasta qué punto y en qué medida sea censurable no es fácil de determinar por la razón, porque no lo es ninguna de las cosas que se perciben. Tales cosas son individuales y el criterio reside en la percepción. (Aristóteles, 1981:1109b)

Varios particularistas contemporáneos se identifican con Aristóteles;<sup>1</sup> sin embargo, nunca hubo acuerdo en considerar al filósofo como particularista o universalista; los estoicos y los pensadores cristianos hicieron de su ética una lectura universalista a partir de la cual sostuvieron que el bien supremo no se conquista solo por el carácter virtuoso sino que está orientado también por la ley natural (O'Neill, 1996:14). En cuanto al universalismo contemporáneo, tanto en su vertiente deontológica heredera de Kant como en la consecuencialista —en especial el utilitarismo—, niega que los principios éticos aceptables hoy posean una base metafísica, aunque es verdad que esto ha sido puesto en duda.<sup>2</sup>

Más difícil resulta reunir en un grupo homogéneo a los particularistas. Como bien señala Guillermo Lariguet (2018:179), bajo el concepto de particularismo se ha pretendido englobar posiciones bastante disímiles: las éticas de la virtud, cuyos representantes contemporáneos más influyentes son Philippa Foot (1994) y Asdrail MacIntyre (2001); las éticas del cuidado, prove-

1 Véase, por ejemplo: Taylor (1985), Williams (1985), McIntyre (2001).

2 Tal es el caso de Sandel en *Liberalism and the Limits of Justice* (1982), obra en la que argumenta que la teoría de la justicia de Rawls no puede desembarazarse del lastre metafísico kantiano.

nientes del pensamiento feminista, que ponen el acento en las relaciones personales más que en las virtudes del carácter, como es el caso de la humeana Annette Baier (1985); algunos pensadores comunitaristas identificados con la vertiente anti-ilustrada, como Sandel y MacIntyre; filósofos narrativistas (Ricoeur, 1991); también quienes valoran la literatura por considerarla una auxiliar de la ética (Diamond (1988:552–564), Charon (1997:91–112)). Lo que comparte este variopinto espectro es su cuestionamiento al rol que los principios juegan en la vida moral; privilegian en cambio la aprehensión contextual y la función de la percepción moral para resolver adecuadamente las situaciones. Este trabajo se enfoca en esta cuestión.

Ahora bien, no todos los particularistas acuerdan en la manera de valorar los principios; algunos niegan su existencia (particularismo radical o fuerte), otros no la descartan al reconocer que los principios cumplen una función, pero no la consideran el elemento prioritario en la formulación y justificación de los juicios (particularismo amplio o débil). Estos últimos suelen adherir a una posición intuicionista.

## Universalismo

Como ya se señaló, las dos tradiciones universalistas de las que abrevan las teorías que hoy defienden la existencia de principios morales son el deontologismo de corte kantiano y el utilitarismo. No me referiré a sus conocidas diferencias sino a los rasgos que comparten y que ameritan identificarlas como universalistas.

En primer lugar, la distinción entre moral positiva y moral crítica. La primera es contingente, cambiante, heterogénea y particularista. Frente a ella, la moral crítica se constituye con criterios argumentativos que funcionan como test de justificación de la moral positiva, lo que permite diferenciar entre facticidad y validez (Nino, 1984:79–82; Guariglia, 1996:11–18; Hare, 1981:44–65).<sup>3</sup>

En segundo lugar, la imparcialidad, que prescribe que los juicios morales no deben estar distorsionados por prejuicios o favoritismos, condición que, desde la modernidad, se sustenta en la igualdad de las personas. El utilitaris-

---

<sup>3</sup> Es ilustrativa la relación entre moral positiva y moral crítica que proporciona Hare en el capítulo 3 de *Moral Thinking* (1981). Allí recurre a las figuras del proletario y del arcángel para representar el pensamiento moral común y el crítico. El proletario razona mediante intuiciones y principios *prima facie*; mientras que el arcángel no requiere de intuiciones porque es enteramente racional. Como el pensamiento crítico, a diferencia del común, puede justificarse a sí mismo, tiene prioridad epistemológica, de modo que los principios *prima facie* deben testarse con los principios del pensamiento del arcángel.

mo la expresa en el principio enunciado por Bentham: «Cada cual cuenta por uno y por nadie más que uno», recogido en las diferentes variantes posteriores de la teoría; es decir, se deben tener en cuenta tanto los propios intereses como los de los otros involucrados por la acción, a fin de alcanzar la posición de un observador imparcial.

El deontologismo contemporáneo prescribe la imparcialidad abrevando del imperativo categórico de Kant derivado del análisis del concepto de obligación. En efecto, cuando el filósofo se refiere a la contravención de un deber señala: «Si ahora atendemos a nosotros mismos en todos los casos en que contravenimos un deber, hallaremos que realmente no queremos que nuestra máxima deba ser ley universal, pero nos tomamos la libertad de hacer una excepción para nosotros mismos» (Kant, 1967:77). Esta idea inspiró el requisito de imparcialidad que está representado por el velo de la ignorancia en el diseño de la posición original de la teoría de la justicia de Rawls (1999:118-122); también la encontramos en la búsqueda del punto de vista *desde ningún lugar* preconizado por Thomas Nagel (1986:80-81) a fin de lograr un juicio moral correcto; el mismo implica colocarse en un punto de vista no personal sino más general que cualquier persona pensante puede comprender. Por su parte Hare pretende derivar la imparcialidad del significado prescriptivo y universalizable del término *deber*, el que conduce a un procedimiento de decisión mediante el cual, «cuando yo prescribo una acción que afecta a otro de cierto modo, dado que estoy prescribiendo universalmente para todas las situaciones similares a la mía, debo ocupar cada uno de los roles alcanzados por dicha acción»<sup>4</sup> (Hare, 1989:215). Como último ejemplo mencionemos la ética dialógica de Jürgen Habermas, en la que es el principio de universalización (U) el que provee el criterio de imparcialidad:

cada norma válida habrá de satisfacer la condición de que las consecuencias y los efectos secundarios que se siguen de su acatamiento *general* para la satisfacción de los intereses de *cada* persona (presumiblemente) puedan resultar aceptados por *todos* los afectados (así como preferidos a los efectos de las posibilidades sustitutivas de regulación). (1991:86)

En tercer lugar, la importancia dada a la razón práctica. La vertiente deontológica del universalismo contemporáneo se caracterizó por su concepción procedimental de la racionalidad práctica. En efecto, en una era posmetafísica, esta ya no puede fundarse en un sujeto trascendental como el kantiano, sino que intentará centrarse en la reconstrucción de mecanismos proce-

---

4 Trad. propia.

dimentales mediante los cuales la razón práctica ya no es un *factum* sino que se justifica a sí misma mediante mecanismos que servirán para fundamentar normas o principios básicos. Esto se observa tanto en propuestas netamente formalistas como la de Habermas (quien propone una reformulación de la razón práctica entendida como una racionalidad práctico-comunicativa que está incorporada a los procesos de diálogo y a los procedimientos argumentativos) y en la tesis formal semántica de Hare,<sup>5</sup> como en las constructivistas sustantivas de Rawls<sup>6</sup> o de Carlos Nino (que justifica sus tres principios morales conectándolos con las características formales del discurso moral).<sup>7</sup>

### **Características de los principios**

La función de los principios morales (sea cual fuere su estatus epistemológico u ontológico) es contribuir a la resolución de los conflictos; esto hace necesarias condiciones de aceptabilidad que son provistas por ciertas características formales (atribuibles tanto a las versiones deontológicas como al utilitarismo) a saber: generalidad (a), consistencia (b), abstracción (c) y universalidad (d):

a) La enunciación de los principios debe ser general, dado que si estuvieran enunciados mediante términos que podrían ser reconocidos por nombres propios o descripciones definidas no cumplirían su función (Rawls, 1999:157); más específicamente: los principios morales y las normas que se infieran de ellos deben ajustarse a dos condiciones: las propiedades que definen clases de sujetos así como los verbos que definen acciones (permitidas o prohibidas) no deben contener términos individuales (Guariglia, 1996:93).

---

5 Una dificultad reiteradamente señalada en la tesis metaética de Hare es su pretensión de derivar del principio de universalización exclusivamente formal un contenido sustantivo: «Si yo estoy tratando de dar un peso igual a los intereses equiparados de todas las partes en una situación, debo considerar un beneficio o perjuicio para cada una de las partes como poseyendo un valor o un disvalor igual a un mismo beneficio o perjuicio para cualquiera de las otras. Esto parece significar que voy a promover del modo más intenso los intereses de las partes; al conceder un peso igual a todos ellos, si maximizo los beneficios totales sobre toda la población aplico el principio clásico de la utilidad» (Hare, 1989:215. Trad. propia). Una crítica detallada a este paso de importación —y no de derivación— del principio utilitarista se encuentra en Gorr (1985:115–138).

6 La posición original puede leerse como una escenificación de las dos capacidades de la razón práctica explicitadas por Kant: el aspecto prudencial y el autolegislatante. El propio Rawls otorga esta interpretación en «El constructivismo kantiano en la teoría moral» (Rawls, 1986:187–170).

7 Nino propone un equilibrio reflexivo más amplio que el empleado por Rawls en la justificación de la posición original llevándolo a tres términos: las convicciones intuitivas, los principios sustantivos generales y las reglas formales del discurso moral, a saber: la publicidad, la universalidad, la generalidad y la imparcialidad (Nino, 1984:88–93).



b) La consistencia fue claramente formulada por Kant como la exigencia que deben cumplir las máximas: «Algunas acciones están de tal modo constituidas que su máxima no puede, sin contradicción, ser siquiera pensada como ley natural universal» (Kant, 1967:76).<sup>8</sup> En la licencia que uno se toma para favorecer sus propios intereses se hará evidente la contradicción que ocurre al universalizar una máxima que lesiona la imparcialidad. Hare la retoma en su tesis prescriptivista: los términos del lenguaje moral tienen un significado prescriptivo y universalizable, dos propiedades formales y lógicas de «deber», verbo característico del lenguaje moral: cualquier juicio moral enunciado por el hablante en determinadas circunstancias lo obliga a sostenerlo en toda otra circunstancia semejante en sus aspectos relevantes por el mero carácter lógico de los términos que utiliza. Por tanto, una persona que dice «yo debo actuar de una cierta manera pero nadie más debe actuar de esa manera en circunstancias similares en sus aspectos relevantes» se está contradiciendo a sí mismo<sup>9</sup> (Hare, 1963:32). La consistencia es, asimismo, la propiedad reclamada tanto por Apel como por Habermas en la reconstrucción de los presupuestos argumentativos orientados a fundamentar la ética discursiva. En su polémica con el falibilista Albert, Apel pretende mostrar que los presupuestos pragmáticos del lenguaje no pueden negarse sin autocontradicción performativa (Apel, 1975:140–173). Habermas hace de dicha contradicción la noción central para respaldar el principio (U):

El teórico moral intentará ahora adoptar la función del escéptico a fin de comprobar si el rechazo de un principio moral propuesto incurre en una realización contradictoria con los presupuestos inevitables de la argumentación moral. De este modo indirecto, el teórico moral puede hacer ver al escéptico que éste, al iniciar una argumentación concreta con el objetivo de refutar el cognitivism ético, tiene que hacer inevitablemente presupuestos argumentativos, cuyo contenido propositivo contradice su objeción (...) en consecuencia, la exigida fundamentación del principio moral recomendado podría adoptar la forma de que toda discusión, en cualquier contexto que se dé, descansa sobre presupuestos pragmáticos de cuyo contenido propositivo puede deducirse el postulado de universalidad U. (Habermas, 1991:104)

---

8 Recuérdese que Kant alude a dos tipos de contradicciones en las que se incurre cuando se universaliza una máxima que viola el imperativo categórico: una de ellas —la transcripta en la cita— es de carácter exclusivamente lógico y ocurre cuando se pretende universalizar una máxima que viola los deberes perfectos; la otra es una contradicción de la voluntad, que aparece cuando la máxima viola deberes imperfectos (Kant, 1967:74–77).

9 Trad. propia.

c) Por mor de su universalidad los principios son necesariamente abstractos, debiendo dejar de lado propiedades diferenciables tanto de los agentes como de los contextos; la abstracción se relaciona también con la base informacional que la teoría en cuestión considera moralmente importante. En el caso del utilitarismo, la base informacional es brindada exclusivamente por la utilidad, sea esta entendida como placer, satisfacción de deseos o preferencias, o demanda de utilidades. El principio de maximización utilitarista hace abstracción de cualquier otra característica que, desde otra perspectiva, podría considerarse éticamente relevante, como por ejemplo los derechos, las libertades o el carácter. El imperativo categórico de Kant exige separar lo universal de lo particular para atenerse únicamente a la forma de la ley, lo que implica no solo dejar de lado las inclinaciones sino todo elemento proveniente de lo contingente, lo que le valió la lapidaria crítica de Hegel al vacío del puro deber (Hegel, 2004:135). El deontologismo contemporáneo —sobre todo en sus variantes sustantivas— procuró hacerse cargo de esta crítica. Por ejemplo, los tres principios que Carlos Nino pretende inferir de su metaética constructivista —el de inviolabilidad de la persona, el de autonomía y el de dignidad—, si bien recogen los conceptos normativos claves de la ética kantiana, introducen explícitamente como base informacional relevante el valor de la libertad individual, lo que permite derivar de sus principios una serie de derechos básicos (Nino, 1984:109–186). Rawls hace algo similar con su primer principio de justicia, mientras que en el segundo principio añade aspectos sociales y económicos como base informacional significativa (Rawls, 1999:52–87).

d) El alcance de la universalidad depende de qué clase de principios se trate. El principio de maximización del utilitarismo posee claramente un alcance plenamente universal. Lo mismo ocurre con los principios formales, como el imperativo categórico o el principio de universalización de Habermas, así como con el test de universalidad que Hare deriva del carácter formal del concepto de deber (Hare, 1999:21–26). Más complejo es determinar el alcance de los principios de justicia de Rawls. Al respecto, su autor parece haber variado de parecer a lo largo del tiempo. En la justificación final que brinda en *Una teoría de la justicia* sostiene que observar nuestro lugar en la sociedad desde la perspectiva de la posición original:

es observarlo *sub specie aeternitatis*: es contemplar la situación humana, no solo desde todos los puntos de vista sociales, sino también desde todos los puntos de vista temporales. La perspectiva de la eternidad no es una perspectiva desde cierto lugar más allá del mundo, ni el punto de vista de un ser trascendente; más bien

se trata de una cierta forma de pensamiento y de sentimiento que las personas racionales pueden adoptar al interior del mundo.<sup>10</sup> (1999:514)

Sin embargo, luego de la revisión de su propia concepción que emprende a partir de la publicación de «Justice as Fairness: Political not Metaphysics» (1985:223–251) restringe el alcance de los principios a las democracias occidentales desarrolladas. Por su parte, Carlos Nino supone que el alcance plenamente universal de sus principios puede justificarse a partir de las condiciones formales de su constructivismo discursivo, condiciones que los hacen aceptables para cualquier potencial participante un discurso moral (Nino, 1981:91).

Sea como fuere, es importante dejar en claro (sobre todo teniendo en cuenta la impugnación que de la universalidad de los principios hace el particularismo) que no todos requieren suponer un alcance universal. Los principios y normas de ellos derivadas tienen determinado grado de universalidad establecido por sus términos que circunscriben clases de individuos o acciones. Esto lo expone con claridad Onora O'Neill en su propuesta de un universalismo kantiano de corte constructivista. La filósofa establece una diferencia entre forma y alcance: algunos principios poseen un alcance plenamente universal y otros relativo a un dominio de agentes; por ejemplo: «Los hijos deben asistir a los padres ancianos», su dominio se restringirá a los hijos que tengan padres ancianos, por otra parte, el significado de asistencia y de ancianidad presentará variabilidades según las características de la sociedad a la que se aplique (O'Neill, 1996:75–77).

## Réplicas particularistas

Estas tesis del universalismo han sido criticadas por un amplio espectro de posiciones de la filosofía práctica: por las versiones contemporáneas de las éticas de la virtud, que, siguiendo a Aristóteles, reclaman un juicio práctico situacional que además tenga en cuenta las cualidades morales del agente y no haga foco exclusivo en los resultados de las acciones o en un conjunto de deberes o derechos; por algunas feministas, en especial las que se dedican a las éticas del cuidado, corriente inspirada en el influyente libro de Carol Gilligan *In a Different Voice* (1982);<sup>11</sup> por lo que en su momento se conoció como co-

---

10 Trad. propia.

11 De acuerdo con el modelo de desarrollo del razonamiento moral propuesto por Lawrence Kohlberg siguiendo a J. Piaget, las mujeres no alcanzan la etapa superior que conlleva el punto de vista de la imparcialidad quedándose en una etapa intermedia en las que identifican la moral con la amabilidad y tienen problemas para utilizar principios abstractos de justicia. Gilligan objeta el

munitarismo, sobre todo en su vertiente antiilustrada (MacIntyre, 2001; Sandel, 1982) e incluso por los bioeticistas que no encuentran en el universalismo orientaciones útiles para la práctica de los profesionales de la salud y que también reclaman un razonamiento práctico contextual que suelen privilegiar las éticas del cuidado.

A pesar de las divergencias que presenta este variado espectro coincide en sus críticas al universalismo ético en las siguientes cuestiones: en primer lugar, el rol dominante y muchas veces excluyente de la razón en la elaboración del juicio moral correcto, así como también la concepción de un modelo idealizado de dicha facultad que supone agentes abstractos cuyas capacidades de cálculo, de libre elección de un plan de vida, de independencia mutua, no son satisfechas por ningún ser humano, lo que significa un completo divorcio de la realidad que atenta contra la aplicabilidad de la teoría.<sup>12</sup> En segundo lugar, un agente moral empobrecido y abstraído de las disposiciones y cualidades que poseemos los seres humanos: por parte del deontologismo, el agente es reducido a la capacidad de aplicar mecánicamente reglas y principios; en cuanto al utilitarismo, se limita a la búsqueda de placer o satisfacción de intereses, lo que conlleva una racionalidad práctica que solo delibera sobre los medios. En tercer lugar, el desconocimiento del papel que juegan las emociones en la elaboración del juicio moral y con ello el desentendimiento de la sensibilidad moral, facultad imprescindible para captar los variados matices que presentan las distintas situaciones en las que debemos actuar. Esta falencia se debe a la incorrecta separación entre razón y emociones amparada en el prejuicio que solo mediante la razón podemos lograr un juicio imparcial.<sup>13</sup> En cuarto lugar, la inutilidad que, debido a su abstracción, poseen los principios para ser eficaces guías de la acción.

---

modelo porque lo encuentra sesgado. Tomando como referencia entrevistas realizadas a varones y mujeres concluyó que utilizan distintas estrategias de razonamiento y dan importancia a cosas distintas cuando resuelven problemas morales. Llamó a las perspectivas empleadas por los varones «de la justicia» y a las mujeres «del cuidado».

12 Objeciones de este tenor se encuentran entre otros, en Sandel (1982); MacIntyre (2001); Pateman (1988); Nicholson, 1985. Una interesante distinción entre abstracción e idealización con la que se propone superar las críticas a las teorías idealizadas la realiza Onora O'Neill (1996).

13 En los últimos años esta tajante separación entre razón y pasiones ha sido ampliamente cuestionada incluso por filósofos que se alinean con el universalismo ético, como por ejemplo Nussbaum (1990, 1995, 2001). Un punto de inflexión en este cambio de perspectiva lo constituye el libro Antonio Damasio *El error de Descartes* (1997). En él, el neurocientífico portugués muestra el rol determinante que juegan las emociones en el razonamiento moral.

## Particularismo

Aunque no todas las críticas referidas en la sección anterior pertenecen a filósofos particularistas, todas ellas son compartidas por estos. En términos generales, el particularismo sostiene que la raíz del juicio ético no se encuentra exclusivamente en la capacidad racional sino también —y especialmente— en la percepción de las situaciones particulares cuyas múltiples aristas y especificidades son captadas por la sensibilidad moral de cada agente; asimismo consideran que el juicio moral está condicionado por las relaciones personales y por los valores socialmente compartidos, elementos todos imposibles de ser capturados por principios generales. Además, ciertos autores, como por ejemplo Jonathan Dancy, endilgan al universalismo su incapacidad para dar cuenta del conflicto moral.

Sin embargo, como hemos adelantado al comienzo, no todos los particularistas comparten la misma posición respecto de los principios. Como también señalamos, el desacuerdo en este aspecto ha llevado a agruparlos en dos tipos: radicales o fuertes y amplios o débiles. Comenzaremos por considerar a los primeros.

### Particularistas radicales

Los particularistas radicales, tales como John MacDowell (1979:331-350), el nombrado Dancy (2013) o Peter Winch (1972) niegan cualquier función a los principios morales; Dancy, incluso, llama la atención sobre lo dañinos que pueden estos resultar como guías orientativas ya que son rígidos y ciegos a las diferencias por lo que demandan un tratamiento inflexible y uniforme.

MacDowell sostiene que la ética debe centrarse en la virtud, a la que define como «la habilidad de reconocer los requerimientos que una situación impone a la conducta» (2002:53). Entendidas como capacidades perceptivas, las virtudes forman parte de nuestra sensibilidad moral a la que considera una capacidad cognitiva que abre nuestra percepción al reino de los valores. Además, siguiendo una línea wittgensteniana, este autor otorga suma importancia al rol que desempeñan en la educación moral las prácticas compartidas comunitariamente; son ellas las que nos permiten acceder y comprender los valores que compartimos con nuestra comunidad.

Más radical que McDowell en su particularismo, Dancy parte de mostrar los defectos de las teorías que apelan a principios (sean estos de carácter absoluto, como el imperativo categórico kantiano o el principio utilitarista de mayor felicidad, o de índole *prima facie* —como el intuicionismo de Ross y

sus seguidores— para defender que el juicio y razonamiento moral no consisten en la aplicación de los principios a los casos. Si se considera que los principios son absolutos, el resultado de su aplicación puede conducir a una catástrofe moral. Al respecto, el trillado ejemplo que siempre viene a la mente para criticar la ética de Kant es la defensa de no mentir en ninguna circunstancia. El filósofo la esgrimió proponiendo el caso de un asesino que concurre a mi casa preguntando si allí se oculta un amigo mío al que di refugio y al que pretende asesinar (Kant, 1987:61–68). Respecto del utilitarismo, la crítica se enfoca en la única característica que considera moralmente relevante, lo que acarrea un reduccionismo injustificable de la variedad de conceptos morales que empleamos para tomar las decisiones. Por otra parte, sostiene Dancy que ninguna de las dos posiciones da cuenta del conflicto moral porque, dado el carácter definitivo de sus principios, ellos siempre aportan la solución correcta y todo conflicto resulta ser solo aparente. Más plausible encuentra la perspectiva intuicionista, ya que esta admite una pluralidad de aspectos relevantes que puede presentar una acción. En lo que no acuerda es en el modo en que el intuicionismo pretende resolver los conflictos. En efecto, Ross parece creer que existen principios morales porque le resulta obvio que si un rasgo es moralmente relevante en un caso debe serlo en cualquier otro en que dicho rasgo se presente, de manera que, si en determinada circunstancia dejamos de lado ese rasgo y privilegiamos otro, debemos dar razones valederas de nuestra opción, luego de un cuidadoso balance de los deberes *prima facie* que están en juego (Ross, 1972:153–161). Sin embargo, argumenta Dancy, que un rasgo sea moralmente relevante en un caso puede estar influido por circunstancias que son específicas de él y que determinan su relevancia. Supongamos que prometí a un amigo acompañarlo a comprar un equipo de música el sábado por la mañana, pero se presenta mi anciana vecina pidiéndome que le ayude a cortar la enredadera que está estrangulando a su joven abedul. El que yo decida cumplir la promesa con mi amigo y no ayudar a mi vecina en ese momento, aunque sí en el día de mañana, o que decida hacer lo contrario, dependerá de circunstancias tales como si a mi amigo le resulta problemático ir solo a hacer esa compra porque no entiende de equipos, o si no lo perjudica que lo acompañe en otro momento, o de que ya le haya prometido acompañarlo la semana anterior, pero no pude hacerlo por razones de trabajo, de que sé que mi anciana vecina se angustiará si no le ayudo ahora o le dará lo mismo postergar la tarea, o de que pienso que prestarle auxilio en este momento es un modo de retribuirle los ricos dulces caseros que suele traerme; en fin, la obligación de cumplir o no cumplir con mi promesa dependerá de mi percepción de todas estas eventualidades. El desacuerdo fundamental de Dancy con Ross es que para estos rasgos morales son deberes que no tie-

nen una relevancia variable por más que en determinado caso esté justificado que un deber sea sobrepasado por otro; en pocas palabras, que esté justificado que en cierto acto cumplamos con el deber de solidaridad en detrimento del de justicia, significa que realizaremos un acto bueno, pero injusto. Esto es lo que Dancy no admite, rechazando así la invariancia de razones ya que, a su modo de ver, estas —tanto las prácticas como las teóricas— son siempre contextuales. Cabe preguntarse entonces por el criterio de corrección del juicio: si no hay principios de ninguna clase ¿qué hace la diferencia entre una buena y una mala razón? La propuesta parece deslizarse hacia un subjetivismo ético, aunque Dancy niega esta posibilidad argumentando que hay un aprendizaje moral que realizamos mediante la experiencia; el poseedor de un juicio moral experto conoce el propósito práctico de los conceptos morales; por ejemplo, habiendo llegado a captar, debido a la práctica acumulada, el propósito práctico del concepto de crueldad es posible que en una situación nueva pueda distinguirse en ella el rasgo «crueldad». La experiencia en situaciones que guardan alguna semejanza nos anoticia sobre cuáles son la clase de cosas a las que debemos prestar atención, cuál es la relevancia que puede tener determinado aspecto, de ello resulta que, cuando se nos presente un hecho con algunas características similares a otros anteriores podremos formular un juicio informado, pero esto no implica que apliquemos al nuevo caso algún principio que hemos extraído de otros anteriores (Dancy, 2013). Surge entonces preguntarse por el criterio que valida las buenas razones. Dancy no lo deja claro, pero, a mi entender hay dos opciones: o bien la subjetivista ya aludida: el criterio lo proporciona mi propia experiencia; o bien la relativista: la validación la proveen los valores aceptados por la moral social, las intuiciones morales socialmente compartidas; sin embargo, en esta alternativa, no resulta impropio argumentar que dichas intuiciones tienen la forma de principios: «No se debe robar», «no se debe mentir» y demás, opción que nos conduce a los particularistas débiles.

### Particularistas débiles

Los que se identifican con un particularismo amplio o débil (McNaughton, 2000:256-275; Little, 2000:276-304; Rawling, 2000:256-275; Audi, 2006) critican, igual que los radicales, el punto de vista según el cual lo que debe hacerse en cada caso pueda determinarse mecánicamente aplicando algún principio, pero, como anticipamos, defienden la función de principios generales; estos nos ayudan a captar los rasgos morales que ofrecen las situaciones, aunque los consideran insuficientes para realizar una correcta evaluación mo-

ral, hecho que no depende de la aplicación de los principios sino de la capacidad perceptiva.

Esta corriente suele estar influida por el intuicionismo de David Ross, de modo que, en primer lugar, nos referiremos sucintamente a sus ideas principales.

Catalogar a Ross como universalista o particularista resulta problemático —cosa que generalmente suele ocurrir con toda pretensión de clasificación de una teoría, siempre a riesgo de caer en reduccionismos—. Ross comparte algunas ideas con las que hemos caracterizado al universalismo y tiene una innegable influencia de Kant. Como este, su preocupación se concentra en el concepto de deber, pero, consciente de las dificultades que presenta el absolutismo kantiano, propone superarlas con su teoría de los deberes *prima facie*. Ross supone la existencia de un número indefinido de principios a los que accedemos por inducción intuitiva a partir de su reconocimiento en los casos particulares. Dado que los asemeja a los axiomas de la matemática, los considera verdaderos y universales pero no en sentido absoluto: no siempre es nuestro deber cumplir una promesa o no mentir; un principio no puede ser verdadero si su aplicación causa un mal. Dejando de lado la problemática cuestión de la verdad que ofrece su tesis, lo que importa tener en cuenta aquí es que, para Ross, a diferencia de Kant, las acciones que ofrecen algún tipo de conflicto poseen más de un rasgo moral relevante (cosa crucial para el particularismo), es decir, presentan más de un deber *prima facie*. ¿Cómo decidir cuál de todos los deberes *prima facie* involucrados se convierte en deber en sentido pleno para la ocasión? Según el autor un acto es correcto cuando de todos los actos posibles para el agente en esas circunstancias será aquel cuya corrección *prima facie* en aquellos aspectos en que es correcto en mayor medida supera su incorrección en aquellos aspectos en los que es incorrecto *prima facie*. Dado que no existe una ordenación de los principios no hay modo de saber a ciencia cierta si la respuesta a que se ha arribado luego de deliberar es la correcta —o incluso, si tal respuesta existe, cosa que Ross deja indeterminado—. La solución aristotélica que ofrece es recurrir a la figura de la persona experimentada. En definitiva, es quien posee sabiduría moral —«la gente cuyo juicio hemos aprendido a respetar más en lo que concierne a cuestiones éticas» (Ross, 1972:166)— quien decidirá de modo adecuado y dará el criterio de corrección de los juicios. Por último, podemos destacar que los principios universales, en su carácter de obligaciones *prima facie*, funcionan como límites de las acciones que no son justificables ya que, dada una situación en la que colisionan dos o más principios, es necesario argumentar por qué se elige uno y se desechan los otros.

David Mc Naughton y Piers Rawling (2000:256–275) elaboran una versión del intuicionismo rossiano en términos de razones justificatorias del juicio.



Contra el particularismo radical de Dancy y especialmente contra las críticas de este a Ross, los autores sostienen que la mejor manera de interpretar su teoría es mostrar que ofrece razones primarias y secundarias para justificar los juicios; las primeras poseen un valor invariable (bueno o malo, correcto o incorrecto, etc.) mientras que los valores de las secundarias son variables, de modo que tienen fuerza moral solo en la medida en que su presencia en contextos particulares borra la fuerza de la razón primaria. Volvamos al ejemplo kantiano de la obligación de no mentir. Decir la verdad es un deber *prima facie* porque normalmente mentir causa algún tipo de daño; sin embargo, decir la verdad al asesino conduciría al homicidio del amigo, razón secundaria que, en este caso, sobrepasa a la razón primaria que acompaña al deber de no mentir; sin embargo es incorrecto afirmar, como hace Dancy, que el valor moral positivo de no mentir cambie a negativo o a neutro; solo significa que el valor moral de la veracidad fue sobrepasado en este caso por el de la solidaridad o beneficencia, dado que una situación puede tener varios rasgos moralmente relevantes; esto conduce a McNaughton y Rawling a concluir que hay algo del valor moral de la acción que se pierde si nos colocamos en la posición de mentir por buenas razones secundarias. No entender que la justicia, o la veracidad, u otros conceptos densos tienen un valor moral positivo invariable es no comprender cómo funciona el juicio moral; de allí que los principios sean útiles para ayudarnos a captar los rasgos morales importantes cuando se presenta un caso complejo, en el que no es sencillo determinar a primera vista qué es lo correcto.<sup>14</sup>

### **Réplicas universalistas**

Los particularistas radicales sostienen que los principios morales del tipo que fuere son superfluos dado que no sirven para la función principal que todo universalista les endilga, esto es, guiar las conductas; sin embargo, esta postura nos deja ante el problema no menor de la injustificabilidad de los juicios. En el caso de Dancy, su rechazo a reconocer la invariabilidad de los aspectos morales de cada acción deja librado al criterio del agente, enriquecido por su experiencia, la correcta resolución que el acto demanda. Pero este es un criterio de corrección completamente insuficiente; ni siquiera recurre al juicio del hombre prudente, como Aristóteles o Ross —sin duda discutible pero que al menos posee alguna pretensión de objetividad—. La teoría de McDowell tiene el problema que presentan las éticas de la virtud que desechan la apela-

---

14 Para una interesante defensa del intuicionismo ver Lariguet (2017).

ción a principios o reglas. La sinceridad y la consideración son dos virtudes: supongamos una situación en que un amigo le da de leer a otro su primera novela. A aquel le parece malísima, ¿cómo resuelve el conflicto? ¿Debe ser sincero o amable?<sup>15</sup> Además, su posición tiene el problema que suelen presentar los particularismos: los valores morales solo pueden ser comprendidos por las prácticas compartidas por las comunidades, de modo que el criterio de corrección es exclusivamente interno, lo que ofrece una dificultad obvia: ¿qué debemos entender hoy por comunidad de prácticas compartidas, dada la heterogeneidad y el pluralismo de las sociedades contemporáneas?: ¿la de las minorías sexuales?; ¿las de los pueblos originarios?; ¿la de los negros?; ¿la de los discapacitados?; ¿la de los hombres blancos?; ¿la de los académicos? Y un largo etcétera. Además, atendiendo a la dimensión cosmopolita, no aporta criterio alguno a los acuciantes problemas de justicia internacional, tales como la discriminación, la opresión, la pobreza, dejando al desnudo el flanco más débil del relativismo ético.

En lo que respecta al particularismo amplio, dado que pretende ser una puesta a punto del intuicionismo, presenta los mismos problemas que este. Uno de ellos es el que expuso Rawls (1999:30–40): los principios no presentan un orden jerárquico que oriente las elecciones en caso de conflicto; el otro se relaciona con el carácter sui generis de las intuiciones morales y su acceso a ellas. ¿Cómo saber si se tienen intuiciones correctas? Ya recordamos que el criterio de corrección que propone Ross es el del sabio moral (1930:41); pero ¿a quién erigimos como sabio moral?; además, ¿un sabio moral no puede errar? Esta pauta, como bien destaca Hare (1999:99), presenta otra dificultad: podría funcionar en las culturas homogéneas, con una educación moral común, pero se torna mucho más problemático en culturas pluralistas. En definitiva, el criterio proporcionado por el intuicionismo no deja de ser una apelación acrítica a la autoridad.

## **A modo de conclusión**

Algunas de las críticas formuladas por los particularistas a las propuestas universalistas resultan razonables. Es verdad que las teorías canónicas han descuidado al agente moral, sea reduciéndolo a un receptáculo de deseos, sea identificándolo con una racionalidad idealizada; los seres humanos no somos ni actuamos así, y si una teoría está demasiado distanciada de la realidad, presentará fallas en su aplicación. También es cierto que se ha prestado escasa atención a las parti-

---

15 Para un muy claro tratamiento de este aspecto puede consultarse (Rachels, 2007:287).

cularidades de las situaciones y, como consecuencia, no se ha tenido en cuenta la competencia de la sensibilidad moral para lidiar con ellas. Creo, sin embargo, que otras objeciones están desencaminadas. Ninguna de las teorías universalistas interesantes propone una aplicación mecánica de los principios; la función de estos es estructurar y restringir las acciones funcionando como pautas; cierto es que, debido a su necesaria abstracción, muchas veces las dejan indeterminadas —otra frecuente crítica de los particularistas—; pero esta indeterminación no es una carencia sino una consecuencia de entender que los principios no pueden aplicarse en piloto automático.

En cuanto a la crítica según la cual el universalismo soslaya el meollo de la vida moral, el conflicto, hay que tener en cuenta que la función de las teorías es justamente contribuir a su resolución. A mi modo de ver, esto no hay que interpretarlo en el sentido de que las teorías principistas suponen que los conflictos son solo aparentes (como por ejemplo señala Dancy al comentar el intuicionismo de Ross). Lo que pretenden es proveer estándares de justificación aceptables que contribuyan a resolverlos. Los principios no son verdades últimas e inapelables; esto lo dejan en claro varios autores, como es el caso de Rawls al proponer el equilibrio reflexivo como justificación (siempre perfectible) de la posición original, también Habermas, al considerar que el principio de universalización y el principio del discurso se limitan a especificar las condiciones formales que deberían cumplir las normas para poder ser aceptadas por todos los afectados. En este caso no es ocioso tener en cuenta que las prácticas discursivas ocurren cuando se producen conflictos políticos y morales, es decir, cuando las certezas normativas compartidas son puestas en cuestión. Incluso puede acordarse con las críticas al vacío formalismo que, desde Hegel, se hicieron al imperativo categórico kantiano, sin embargo, vale la pena tener en cuenta que se trata de un criterio limitativo destinado a testar la moralidad de las máximas que proporciona razones para abstenerse de violar las obligaciones. Ciertamente es que los principios intuicionistas pueden ser tomados como verdades últimas, pero, aun en este caso, su carácter pluralista impide que se los considere ineludibles.

En lo que respecta a la falta de preocupación del universalismo por temas de psicología moral tales como la sensibilidad, los aspectos perceptivos o el descuido de las virtudes, no es una falencia que presenten todas las posiciones universalistas. Al respecto resulta iluminador el constructivismo universalista de Onora O'Neill. La filósofa, inspirándose en Kant, elabora una compleja teoría de la racionalidad práctica en la que, a diferencia de Rawls, Habermas, Dworkin, o Nagel, suscribe principios abstractos pero no idealizados, con el objetivo de resolver los problemas de reconocimiento por parte de los agentes reales que supone la idealización; además incorpora las virtudes

como condición necesaria para una adecuada justificación de los principios morales (O'Neill, 1996). Otra filósofa que, aun adhiriendo al universalismo —en su caso basado en una teoría del bien humano— presta suma importancia a cuestiones de psicología moral es Martha Nussbaum. En varios de los trabajos en los que defiende la introducción de la literatura para elaborar una teoría ética (1990, 1995) propone, inspirándose en Rawls, el método del equilibrio perceptivo para la deliberación ética que introduce, a diferencia de la propuesta rawlsiana, las emociones que provocan ciertas obras de ficción. Como en el equilibrio reflexivo rawlsiano, se parte de dos polos: uno de ellos provisto por las intuiciones y creencias, al que se añaden las emociones que despiertan en los lectores ciertas obras literarias, y, por el otro, teorías tales como el utilitarismo, el deontologismo, o una teoría del bien. Cuando se llega a un estado de equilibrio perceptivo (coherencia entre las creencias, intuiciones, emociones y la teoría que resulte validada mediante este mecanismo), se habrá alcanzado la verdad (siempre provisoria).

Sin pretensión de analizar aquí ambas propuestas, podemos convenir, en principio, que muestran que no es necesario renunciar al universalismo para dar cuenta de un agente moral de mayor carnadura, capaz de ejercer un razonamiento práctico sensible a los contextos.

Para finalizar, a mi modo de ver, el mayor problema del particularismo —en contraste con la fortaleza del universalismo— es que no distingue entre facticidad y validez, lo que termina convirtiéndolo en conservadurismo moral.

## Referencias bibliográficas

- Apel, K.O. (1975). El problema de la fundamentación última filosófica a la luz de una pragmática trascendental del lenguaje. *Diánoia*, (21), 135.
- Aristóteles. (1981). *Ética a Nicómaco* (trad. de María Arujo y Julián Marías). Centro de estudios constitucionales.
- Audi, R. (2006). Ethical Generality and Moral Judgement. En Dreier, J. (Ed.), *Contemporary Debates in Moral Theory* (pp. 285–304). Blackwell.
- Baier, A. (1985). What do Women want in Moral Theory. En Baier, A. *Moral Prejudices: Essays on Ethics* (pp. 1–17). Harvard University Press.
- Charon, R. (1997). The Ethical Dimensions of Literature. En Hilde Lindenmann, N. (Comp.), *Stories and Their Limits* (pp. 91112). Routledge.
- Damasio, A. (1997). *El error de Descartes* (trad. Pierre Jacomet). Editorial Andrés Bello.
- Dancy, J. (2013). Moral Particularism. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition).  
<https://plato.stanford.edu/archives/fall2013/entries/moral-particularism/>
- Diamond, C. (1988). Argument and Perception: The Role of Literature in Moral Inquiry. *The Journal of Philosophy*, 85, 23–41.
- Foot, Ph. (1994). *Las virtudes y los vicios* (trad. de Claudia Martínez). UNAM.
- Gilligan, C. (1982). *In a Different Voice*. Harvard University Press.
- Gorr, M. (1985). Reason, Impartiality and Utilitarianism. En Potter, N.T. y Timmons, M. (Comps.), *Morality and Universality: Essays on Ethical Universalizability*. D. Reidel.
- Guariglia, O. (1996). *Moralidad. Ética universalista y sujeto moral*. Fondo de Cultura Económica.
- Habermas, J. (1991). *Conciencia moral y acción comunicativa* (trad. de Ramón García Cotarelo). Península.
- Hegel, G.W.F. (2004). *Principios de la filosofía del derecho* (trad. Juan Luis Vermal). Sudamericana.
- Hare, R.M. (1963). *Freedom and Reason*. Oxford University Press.
- Hare, R.M. (1981). *Moral Thinking*. Oxford Clarendon Press.
- Hare, R.M. (1999). *Ordenando la ética* (trad. Joan Vergés Gifra). Ariel.
- Kant, M. (1967). *Fundamentación de la metafísica de las costumbres* (trad. de Manuel García Morente). Espasa–Calpe.
- Kant, M. (1987). Sobre un presunto derecho a la filantropía. En *Teoría y práctica* (trad. Juan Manuel Palacios). Tecnos.
- Lariguet, G. (2018). *Dilemas en la moral, la política y el derecho*. Editorial B de F.
- Little, P.R. (2000). Moral Generalities Revisited. En Hooker, B. y Little, M. (Eds.), *Moral Particularism* (pp. 276304). Clarendon Press.
- Nagel, T. (1986). *The View from Nowhere*. Oxford University Press.
- MacDowell, J. (1979). Virtue and Reason. *The Monist*, 62(3), 331–350.
- MacDowell, J. (2002). *Mind, Value and Reality*. Cambridge University Press.
- McNaughton, D. y Rawling, P. (2000). Unprincipled Ethics. En Hooker, B. y Little, M. (Eds.), *Moral Particularism* (pp. 256–275). Clarendon Press.
- Nicholson, L. (1985). *Feminism and Marx: Integrating Kinship with the Economic*. Praxis international.

- Nino, C.S. (1984). *Ética y derechos humanos*. Paidós.
- MacIntyre, A. (2001). *Tras la virtud* (trad. de Amelia Valcárcel). Crítica.
- Nussbaum, M. (1990). *Love's Knowledge*. Oxford University Press.
- Nussbaum, M. (1995). *Poetic Justice*. Beacon Press.
- Nussbaum, M. (2001). *Upheavals of Thought*. Cambridge University Press.
- O'Neill, O. (1996). *Towards Justice and Virtue*. Cambridge University Press.
- Pateman, C. (1988). *El contrato sexual* (trad. María Luisa Femenías). Anthropos.
- Rawls, J. (1985). Justice as Fairness: Political not Metaphysical. *Philosophical & Public Affairs*, 14, 223–251.
- Rawls, J. (1986). El constructivismo kantiano en la teoría moral. En *Justicia como equidad* (trad. de Miguel Ángel Rodilla; pp. 137–186). Tecnos.
- Rawls, J. (1999). *A Theory of Justice*. Harvard University Press.
- Rachels, J. (2007). *Introducción a la filosofía moral*. Fondo de Cultura Económica.
- Ricoeur, P. (1991). Life in Quest Narrative. En Wood, D. (Ed.), *On Paul Ricoeur Narrative and Interpretation* (pp. 20–33). Routledge.
- Ross, D. (1930). *The right and the good*. Oxford University Press.
- Ross, D. (1972). *Fundamentos de Ética* (trad. de Dionisia Rivero y Andrés Pirk). EUDEBA.
- Sandel, M. (1982). *Liberalism and the Limits of Justice*. Cambridge University Press.
- Taylor, Ch. (1985). *Philosophy and Human Science*. En *Philosophical Papers*. Vol. II. Cambridge University Press.
- Williams, B. (1985). *Ethics and the Limits of Philosophy*. Harvard University Press, Cambridge.
- Winch, P. (1972). *Ethical and Action*. Routledge Library Editions.

## 8. Intuiciones morales

*Guadalupe Mettini*

El trabajo se propone examinar el rol que las intuiciones pueden desempeñar en el razonamiento moral. Ofrecemos para ello, un panorama de la discusión filosófica sobre esta cuestión. Las intuiciones han sido interpretadas en paradigmas diferentes como inclinaciones a aceptar proposiciones como verdaderas, como tendencias a juzgar de cierto modo o como actos de aprehensión intelectual de verdades. Desde algunos puntos de vista las intuiciones desempeñan un rol fundamental en el conocimiento, posibilitando el acceso a entidades abstractas, proporcionando un trasfondo que permite evaluar hipótesis empíricas o facilitando la comprensión de una teoría. El trabajo se estructura en cuatro secciones en las que se presentan críticamente diferentes perspectivas sobre la naturaleza de las intuiciones, se examinan sus funciones en los juicios morales, se propone un análisis de roles de las intuiciones no contemplados por las perspectivas presentadas y se plantea problemas abiertos.

El objetivo de la primera sección es contextualizar la controversia y mostrar cómo son entendidas las intuiciones en la filosofía contemporánea refiriendo algunas de estas perspectivas a sus antecedentes modernos. Presentamos para ello la noción de intuición desde puntos de vista diversos y abocados a problemas filosóficos diferentes. Con ello queremos ilustrar la heterogeneidad de características y funciones que se les han atribuido. En la segunda sección presentamos una exposición esquemática de las perspectivas que han surgido para explicar la naturaleza y el rol de las intuiciones en el razonamiento moral. Realizamos luego una evaluación de dichas perspectivas sobre la base de sus méritos y limitaciones explicativas. Identificamos los que consideramos los dos obstáculos centrales para las perspectivas intuicionistas: elucidar el desacuerdo moral y explicar los actos de disidencia. Mostramos en qué sentido son problemáticos para las tres perspectivas presentadas y proporcionamos salidas posibles a las objeciones que plantean para cada versión. Por último, examinamos un aspecto que creemos no ha sido suficientemente tematizado en las perspectivas visitadas en este trabajo. En general, las intuiciones se asocian a prejuicios o sesgos. Es en virtud de que presentan información parcializada que permiten obtener conclusiones de manera automática. Las diferentes versiones del intuicionismo han intentado mostrar cómo es posible sostener simultáneamente que las intuiciones son fundamentales para el razonamiento moral, pero al mismo tiempo no son rígidas o pueden ser modificadas o in-

cluso desafiadas. En este sentido, las interpretaciones examinadas se han centrado en mostrar que las intuiciones no son reducibles a los prejuicios o que son sensibles a nueva información. Sin asumir un compromiso específico con alguna de las tesis sobre la naturaleza de las intuiciones, sugerimos que estas desempeñan también papeles positivos en el razonamiento moral, importantes para explicar conductas morales y fundamentales en recursos metodológicos que pueden ser empleados para solucionar conflictos entre pares epistémicos. Finalmente sintetizamos las conclusiones de nuestro recorrido y señalamos los problemas abiertos que consideramos requieren de atención para progresar en el debate sobre las intuiciones morales. Entre ellos destacamos: el estatus y rol de las intuiciones en la justificación, la relación entre prejuicios e intuiciones y el papel de las emociones en el razonamiento moral.

### **¿Qué son las intuiciones?**

Las intuiciones pueden ser caracterizadas, de manera general y provisoria, como estados mentales que tienen como objeto una proposición y que son acompañados de un sentimiento de certeza. Si bien la noción de intuición tiene una larga historia en la historia de la filosofía, podemos tomar como punto de partida para abordar su tratamiento en la filosofía contemporánea, algunas de las conceptualizaciones que se elaboraron durante la modernidad. En este contexto, las intuiciones fueron consideradas como fuente de evidencia y en garantía de la verdad para las posiciones racionalistas. En la filosofía cartesiana, las intuiciones se vincularon con el modo de presentación de ciertos contenidos mentales. La claridad y la distinción fueron identificadas como propiedades de las verdades intuitivas.<sup>1</sup> Como medio de acceso al conocimiento, la intuición fue pensada como una vía directa de conocimiento de ideas, una forma de aprehensión intelectual de verdades autoevidentes.<sup>2</sup>

---

1 «En este primer conocimiento no se encuentra más que una percepción clara y distinta de lo que conozco; la cual, en verdad, no sería suficiente para asegurarme de que es verdadera, si alguna vez pudiera suceder que una cosa que llegue a concebir tan clara y tan distintamente se hallara ser falsa. Y por lo tanto, me parece que desde ahora puedo establecer como regla general que todas las cosas que concibamos muy clara y muy distintamente son verdaderas» (Descartes, [1641]2011:178).

2 «Entiendo por intuición no la fluctuante confianza de los sentidos o el juicio falaz de la imaginación que compone arbitrariamente; sino más bien un concepto tan fácil y distinto de la mente pura y atenta, que no nos quede duda alguna de aquello que entendemos; o, lo que es lo mismo, un concepto no dudoso de la mente pura y atenta, nacido de la sola luz de la razón y más cierto que la misma deducción, por ser más simple, aunque ya hicimos notar también que no puede ser realizada indebidamente por el hombre. De esta suerte cada uno puede intuir con el espíritu que



Por el contrario, en perspectivas empiristas como la de Hume, las impresiones son el fundamento del conocimiento. Solo pueden conocerse intuitivamente las relaciones de ideas, cuya verdad no depende de los estados de cosas.<sup>3</sup> En la filosofía kantiana las intuiciones puras fueron interpretadas como condiciones de posibilidad del conocimiento sensible,<sup>4</sup> cuyo contenido está dado por las intuiciones empíricas. La intuición racional como captación inmediata de verdades está vedada también desde este punto de vista, o al menos restringida al conocimiento de principios lógicos.

En la discusión contemporánea el rol de las intuiciones ha sido reevaluado a partir de la relevancia psicológica y epistémica que las ciencias de la conducta les han atribuido. Si bien los enfoques sobre la naturaleza y el rol de las intuiciones se diferencian fundamentalmente por el tipo de compromisos epistémicos y filosóficos que asumen (y aquí podríamos trazar una continuidad entre las distinciones realizadas por los filósofos modernos referidos y las caracterizaciones actuales de la intuición) puede señalarse como un aspecto común la atención a dos de sus rasgos heurísticos, a saber: la instantaneidad (en ocasiones entendida como propia de un proceso no inferencial) y la coerción que ejercen sobre la formulación de juicios. En general las intuiciones están vinculadas a una forma especial de acceso a ciertos contenidos, diferente de la percepción y del razonamiento inferencial. En algunos ámbitos esta forma de acceso al conocimiento ha sido postulada para explicar cómo es posible acceder a entidades con las cuales no es posible establecer una cadena causal. Tal es el caso del conocimiento en matemáticas. Algunos autores sugieren que esta forma de conocimiento es admisible también de los principios que gobiernan el mundo físico, desde este punto de vista las leyes físicas son entidades abstractas que pueden ser captadas intuitivamente del mismo modo que

---

existe, que piensa, que el triángulo está determinado únicamente por tres líneas de la esfera por una sola superficie» (Descartes, [1641]2011:9).

3 «Todos los objetos de la razón e investigación humana pueden, naturalmente dividirse en dos grupos, a saber: relaciones de ideas y cuestiones de hecho, a la primera clase pertenecen las ciencias de la geometría, álgebra y aritmética y, en resumen, toda afirmación que es intuitiva o demostrativamente cierta. Que el cuadrado de la hipotenusa es igual al cuadrado de los dos lados es una proposición que expresa la relación entre estas partes del triángulo, que tres veces cinco es igual a la mitad de treinta expresa una relación entre estos números. Las operaciones de esta clase pueden descubrirse por la mera operación del pensamiento, independientemente de lo que pueda existir en cualquier parte del universo. Aunque jamás hubiera habido un círculo o un triángulo en la naturaleza, las verdaderas demostradas por Euclides conservarían siempre su certeza y su evidencia. Investigación sobre el conocimiento humano» (Hume, [1748]2010:57–58).

4 «Llamo puras (en sentido transcendental) a todas las representaciones en las que no se encuentra nada que pertenezca a la sensación. Según esto, la forma pura de las intuiciones sensibles en general, en la cual todo lo múltiple de los fenómenos es intuido en ciertas relaciones, se encontrará a priori en la mente» (Kant, [1781]2007:88).

los principios matemáticos. Otras perspectivas sugieren que el contacto con el mundo empírico permite a los sujetos acumular información que, si bien no se encuentra formulada proposicionalmente, hace posible el reconocimiento de patrones y con ello facilita la identificación de principios físicos.

Algunas perspectivas filosóficas han asumido que las intuiciones son las formas características del conocimiento matemático. Dentro de esta perspectiva, el punto de vista sustentado por Gödel (1964) interpreta a las intuiciones como la base de justificación de los teoremas matemáticos. Esta lectura es común al grupo de posiciones asociadas con el platonismo matemático (Linnebo, 2017), según las cuales la intuición es una forma de percepción de objetos abstractos. En la misma línea, Brown (2011) entiende que la intuición constituye una forma de conocimiento a priori, un mecanismo no inferencial diferente de la experiencia perceptiva, en virtud de la cual en ocasiones es posible acceder intelectualmente a leyes de la naturaleza (como objetos de la intuición intelectual).

Posiciones comprometidas con la epistemología empirista, han caracterizado a las intuiciones como contenidos cognitivos que los sujetos adquieren de su contacto con el mundo y que de forma inarticulada forman el trasfondo a partir del cual son evaluadas las hipótesis empíricas. Para Mach (1897) por ejemplo, las intuiciones empíricas nos compelen a concebir al mundo físico de cierto modo, no dependen de nuestra voluntad de formarnos juicios sobre los fenómenos y constituyen una forma de conocimiento objetivo. La procedencia de las intuiciones es doble. Por un lado, se forman a partir de las experiencias pasadas del individuo, por otro lado proceden filogenéticamente, de la experiencia acumulada de la especie y como tales configuran la capacidad cognoscitiva del sujeto. A esta interpretación adhieren autores de trabajos recientes como Brendel (2005) y Buzzoni (2009) quienes defienden esta lectura como base de la interpretación de formas de razonamiento especulativo en ciencias fácticas.

Otras perspectivas han examinado el uso de la noción de intuición en la práctica científica. Este concepto es ampliamente usado en el campo de la ciencia física. En este contexto, Tallant (2012) señala que es posible distinguir entre tres acepciones de la noción.<sup>5</sup> Por un lado, un resultado experimental descrito como *intuitivo* es un resultado esperable y, por lo tanto, valorado positivamente. Por otro lado, una teoría *intuitiva* es una teoría que proporciona o permite derivar una explicación de un fenómeno. Dada la ubicuidad de las representaciones visuales en física, la noción de intuición aparece re-

---

5 Tallant (2012) revisa publicaciones recientes de alto impacto en el contexto de la física teórica y experimental y a partir de allí recupera los dos sentidos de intuición mencionados.

gularmente asociada a la capacidad de proporcionar modelos visualizables de la realidad. En síntesis, en el contexto de la física una teoría cuya exposición propicia un entendimiento *intuitivo* es una teoría virtuosa. En el ámbito de la física es habitual la interpretación de acuerdo con la cual las intuiciones se construyen o se suscitan a través de la presentación de visualizaciones y modelos. Este proceso consiste en generar un sentimiento de familiaridad con la representación a partir de la presentación de diagramas y otras representaciones visuales.

En el contexto de la epistemología contemporánea varios autores han reexaminado las caracterizaciones modernas de las intuiciones. La discusión más reciente se ha centrado en el rol que las intuiciones pueden desempeñar en el conocimiento en general pero especialmente en la especulación filosófica. Enfoques como el de Sosa (2007) por ejemplo, rechazan el modelo de la percepción bajo el que tradicionalmente fueron interpretadas las intuiciones y las caracteriza como estados conscientes con contenido, anteriores a las creencias y al conocimiento. Son asimismo estados no factivos, diferenciables de otras actividades como adivinar o tener una corazonada (Sosa, 2007:52). Que son estados no factivos quiere decir que tener una intuición de que «p» no implica que «p» de hecho se dé. En general, para Sosa las intuiciones no derivan de la percepción, la introspección, el testimonio o el razonamiento inferencial, aunque pueden proporcionar una base de justificación para estados doxásticos. Otras perspectivas que asumen que las intuiciones son anteriores a los juicios y a las creencias, las asimilan a un parecer intelectual. Desde este punto de vista: un sujeto tiene la intuición que «p» si le parece que «p», por lo que las intuiciones son estados mentales sui generis no reducibles a otra especie de estado mental. Cuando un sujeto tiene una intuición, la proposición parece ser verdadera para él y por este motivo se siente inclinado a juzgarla de ese modo. Aunque la intuición es en este caso la base del juicio, es diferenciable de este (Boghossian, 2014:368); Bengson (2015), por su parte, afirma que la analogía perceptual es un marco favorable para explicar la naturaleza de las intuiciones ya que estas son presentaciones. Desde esta perspectiva la intuición es un estado mental consciente. La intuición y la experiencia son estados conscientes no factivos, con un contenido que tienen en común con la percepción y con la propiocepción el hecho de presentar al mundo como siendo de cierta manera. Como otros estados presentacionales, se caracterizan además por no haberse formado basados en otros estados mentales, ser graduales (es decir que su calidad puede variar en diferentes circunstancias), ser fundamentalmente no voluntarias, ser convincentes o compulsivas (en tanto fuerza a los sujetos o los dispone con cierta fuerza a asentir sus contenidos) y racionalizar el asentimiento (Bengson, 2015:720-725).

Si bien lo dicho no constituye una reconstrucción exhaustiva del tratamiento de la noción de intuición en la filosofía contemporánea, el recorrido muestra la variedad de rasgos que se les han atribuido a las intuiciones. Las mismas han sido asociadas en el contexto de la filosofía racionalista y en el contexto del realismo matemático a la captación intelectual de objetos abstractos (ideas o entidades matemáticas). En perspectivas empiristas, las intuiciones tienen origen en la experiencia y constituyen un trasfondo de conocimiento accesible directamente a los sujetos a partir del cual es posible reconocer patrones o identificar regularidades. En el contexto de la física, la noción se asocia al sentimiento de familiaridad y accesibilidad que acompaña algunas representaciones. En el debate epistemológico reciente, las intuiciones han sido interpretadas como estados conscientes previos a las creencias, inclinaciones a creer o estados presentacionales similares a las percepciones.

Las diferentes perspectivas presentadas han sido objetadas sobre la base de sus limitaciones explicativas. Las posiciones aprioristas (la concepción moderna y las perspectivas platonistas contemporáneas en matemática y física) han sido criticadas principalmente por los compromisos metafísicos inflacionarios que asumen. Los enfoques empiristas resultan problemáticos porque implican la postulación de una forma de conocimiento sintético accesible a priori. Las perspectivas vinculadas a la epistemología contemporánea en sus diferentes variantes han sido objetadas fundamentalmente sobre la base de la dificultad de establecer criterios de confiabilidad de las intuiciones. La recientemente desarrollada filosofía experimental<sup>6</sup> ha mostrado que existen variaciones significativas entre las intuiciones evocadas por los experimentos mentales filosóficos, dependiendo de la procedencia social, económica, cultural y de género de los sujetos (Stich y Tobia, 2018:428–431). Este hecho representa un obstáculo para postular que las intuiciones pueden justificar creencias. La analogía entre intuición y percepción ha sido criticada debido a las importantes diferencias que subsisten entre estos mecanismos. La principal diferencia entre ambos es la imposibilidad de establecer un vínculo causal con los objetos que son el contenido de las intuiciones. En el caso de la percepción, los objetos físicos se postulan como base de las experiencias sensitivas. La existencia de teorías científicas bien establecidas de las que se derivan pautas de corrección para detectar ilusiones perceptuales y la dificultad para elaborar criterios análogos para las intuiciones son razones de peso para descartar que sean lo suficientemente semejantes. En esta misma línea, puede señalarse que

---

<sup>6</sup> La filosofía experimental designa aquí a la perspectiva filosófica representada por filósofos como Weinberg y Alexander (2006) que emplea métodos empíricos para sustanciar o revisar teorías filosóficas. Entre las estrategias empleadas se encuentran las encuestas y los relevamientos de diferentes respuestas a experimentos mentales.

la intersubjetividad como fundamento para establecer criterios de corrección parece estar ausente en el caso de las intuiciones.

Algunos de los argumentos y objeciones presentadas hasta aquí aparecen también en la discusión acerca del rol de las intuiciones en el razonamiento moral. Como mostraremos en la próxima sección, las intuiciones morales han sido vinculadas a la captación de principios de hechos morales, a inclinaciones a juzgar de cierto modo o a formas especiales de razonamiento. Las diferentes perspectivas coinciden en atribuir a las intuiciones morales un carácter no voluntario, la propiedad de ser instantáneas (aunque no necesariamente no inferenciales) y la característica de ser compulsivas, es decir, de compeler a los sujetos a interpretar casos o a juzgar de cierto modo.

### **Intuiciones y razonamientos morales**

Las intuiciones han tenido un rol importante en el contexto de las teorías éticas y metaéticas contemporáneas. Se ha examinado su papel en los experimentos mentales como el recurso metodológico fundamental en la elucidación de cuestiones éticas y se ha sugerido que las intuiciones desempeñan un papel central en la identificación de los principios generales del razonamiento moral. Las funciones que diferentes perspectivas les han atribuido están vinculadas, naturalmente, a la posición que asumen en términos metafísicos y epistémicos. A continuación, delinearemos tres interpretaciones de las intuiciones morales que consideramos que son representativas del análisis de las mismas en el campo de la ética: el enfoque ligado al intuicionismo moral, el intuicionismo social y la perspectiva que denominamos racionalista.<sup>7</sup> La conceptualización de las intuiciones desde cada enfoque permite delimitar el rol que estas desempeñan en el razonamiento moral.

Las intuiciones morales pueden ser caracterizadas a partir de algunos de los rasgos que hemos recuperado como típicos de las intuiciones en general. Hemos dicho que en general las intuiciones son entendidas como estados mentales, que tienen como contenido una proposición y que están acompañados de un sentimiento de certeza. Su característica principal es que consisten en

---

<sup>7</sup> Las denominaciones que empleamos no coinciden con teorías éticas, sino que se proponen nombrar posiciones que comparten a grandes rasgos una perspectiva epistemológica y una metafísica. El intuicionismo moral se compromete con la existencia de hechos morales y la posibilidad de conocerlos mediante la intuición. Las perspectivas emotivista e intuicionista social comparten un trasfondo empirista y relativista y no se comprometen con la existencia de hechos morales. Por racionalismo entendemos la posición de acuerdo con la cual los juicios morales se formulan a partir de un proceso de razonamiento inferencial aunque desde el punto de vista metafísico no se comprometen necesariamente con la existencia de hechos morales.

formas de acceso directo a estos contenidos. La analogía con la percepción es útil aquí para hacer evidente este rasgo. Así como ciertas cosas pueden aparecer perceptualmente como siendo de cierto modo, por ejemplo, coloreadas o rectas, ciertas proposiciones pueden parecer verdaderas o presentarse a la mente como verdaderas. En el caso de las intuiciones morales éstas se han interpretado como pareceres intelectuales asociados a la manera es que se nos presentan algunas proposiciones (algunas proposiciones nos parecen evidentemente verdaderas) como creencias y como aprehensiones de principios morales. La perspectiva conocida como intuicionismo moral, en sus diferentes formulaciones, adopta una interpretación de este tipo. Desde este punto de vista la intuición es la vía de acceso al conocimiento moral. A través de las intuiciones captamos propiedades o principios morales. Esta perspectiva asume, además, que existen hechos morales en virtud de los cuales ciertos enunciados pueden ser evaluados como verdaderos o falsos. Las perspectivas «emotivistas» en cambio, que encuentran antecedentes en la filosofía empirista de Hume, señalan que el fundamento de los juicios morales no se encuentra en la razón sino en la emoción. Los sujetos son propensos a percibir un acto como correcto o incorrecto, pero los juicios de valor no son ni verdaderos ni falsos porque su fundamento es subjetivo. Los juicios morales expresan emociones y son similares a los juicios estéticos que expresan gusto o disgusto. En esta línea, para el intuicionismo social, los juicios morales son en su mayor parte intuiciones entendidas como sentimientos viscerales, que dan lugar a juicios rápidos y automáticos. Desde este punto de vista, aunque los juicios morales no son producto de razonamientos, los sujetos son capaces de elaborar inferencias que justifican sus posiciones morales. Una interpretación alternativa de las intuiciones entiende que entre razonamiento y captación intuitiva hay una continuidad. Se trata del mismo procedimiento, consistente básicamente en la realización de una inferencia, que es experimentado de manera diferente por los sujetos: las intuiciones se presentan como juicios instantáneos y los argumentos como razonamientos mediatos. A continuación, exponemos con más detalle cada una de estas perspectivas y explicitamos la manera en que consideramos qué se articulan en cada una de ellas las funciones de las intuiciones en el razonamiento moral.

El «intuicionismo moral» no delimita una posición sino un conjunto de enfoques emparentados principalmente por converger en dos principios. Por un lado, los intuicionistas sostienen que las intuiciones son estados mentales que consisten en captar o percibir directamente principios o propiedades morales. Por otro lado, el intuicionismo suscribe un realismo no naturalista, es decir, a la tesis de acuerdo con la cual las propiedades morales no son reducibles a propiedades naturales. Entre las filas del intuicionismo pueden ubi-

carce a autores como Price ([1758]1969), Ross (2001) McDowell (1998), Shafer-Landau (2005), Enoch (2011) o Audi (2013). Aunque existen importantes diferencias entre cada una de las conceptualizaciones elaboradas por estos filósofos, las intuiciones morales se caracterizan dentro de esta perspectiva por ser no inferenciales y tener un fundamento objetivo (con ello, los juicios morales tienen valor de verdad). Por sus rasgos distintivos, las intuiciones morales demandarían una epistemología diferente de la de las intuiciones matemáticas o físicas.

La perspectiva que denominamos «intuicionismo social» está representada por el enfoque de Jonathan Haidt (2001). El punto de partida de Haidt es la distinción entre dos módulos de procesos cognitivos: el sistema intuitivo y el de razonamiento. El primero se compone de procesos automáticos, sin esfuerzo, rápidos, a menudo asociativos, paralelos, afectivos y modulares. Los sujetos solo son conscientes de los resultados de los procesos de este módulo, pero no de los procesos en sí. El segundo módulo se compone de procesos conscientes, esforzados, lentos, que exigen memoria, a veces inferencia y se caracterizan por ser racionales y lineales (Kauppinen, 2018). La intuición moral se puede definir como la aparición repentina en la conciencia de un juicio moral, incluida una valoración afectiva como «bueno-malo», «agrado-disgusto» (Haidt, 2001:818).

El intuicionismo social de Haidt se propone una imagen más «realista» del razonamiento moral, al explicar cómo «de hecho» las personas realizamos juicios morales. Desde este punto de vista, lo que realmente ocurre es que realizamos juicios morales de manera automática e inconsciente a partir de principios socialmente adquiridos. Las intuiciones morales son interpretadas como sentimientos viscerales que determinan los juicios morales. Si bien los sujetos son capaces de razonamiento moral, este no tiene lugar en todos los casos en los que realizamos un juicio, sino que es *post facto* y está destinado a justificar los juicios intuitivos. Esto causa la ilusión de racionalidad en los juicios morales. Los argumentos morales usualmente están destinados a convencer a otros y avalar los propios juicios.

La interpretación que denominamos «racionalista» se distancia de la interpretación que opone intuiciones a razonamientos y propone que intuiciones y juicios están basados en razonamiento producidos por el mismo proceso cognitivo. Desde este punto de vista, intuición y razonamiento son dos fases del mismo proceso que abarca respuestas rápidas e inmediatas y respuestas lentas, conscientes y juicios elaborados. Los juicios morales se producen por el mismo tipo de inferencia que descansa en la inducción la deducción y la abducción. Este enfoque desafía la idea de que las intuiciones morales son rígidas y están dirigidas solo por principios culturalmente aprendidos. La

intuición es experimentada por los sujetos como un proceso inmediato, simple, pasivo y no verbal; mientras que los juicios morales son percibidos como productos de un proceso mediato, complejo, en el que el sujeto es activo y que involucra la formulación de proposiciones y la realización de inferencias. No obstante, el hecho de que los procesos son experimentados como diferentes no garantiza que haya una diferencia en la naturaleza de ambos procesos. Una conclusión específica puede obtenerse sobre la base de un procesamiento inferencial, semánticamente estructurado sin que los sujetos sean conscientes de que la información está siendo procesado (Dellantonio y Job, 2012:243). Este procesamiento puede empezar espontáneamente, ser rápido, y no requerir de una reflexión particular pero aun así depender de operaciones lógicas y semánticas. Para el punto de vista racionalista, la noción de racionalidad no describe solo formas conscientes de procesamiento de la información, ni tampoco incluye ninguna garantía de que la conclusión de la inferencia sea racional en el sentido de la mejor posible.

Presentados de manera esquemática, los tres enfoques constituyen explicaciones plausibles del rol de las intuiciones en el razonamiento moral. No obstante, en cada caso se plantean problemas y limitaciones. Uno de los principales obstáculos del intuicionismo moral es la defensa consistente de la tesis metafísica de acuerdo con la cual existen hechos morales. Asimismo, la opacidad de las intuiciones y el fenómeno del desacuerdo moral arrojan dudas sobre las versiones del intuicionismo de acuerdo con las cuales ciertas proposiciones morales son autoevidentes. Otros problemas para esta perspectiva se siguen de los casos en los cuales aparecen intuiciones morales en conflicto. La ausencia de un criterio para jerarquizar intuiciones ha sido señalada como una razón para desacreditar al intuicionismo. Algunas respuestas posibles a estos problemas consisten en moderar la tesis metafísica, recortándola al no reductivismo, es decir, centrándola en que los hechos o las propiedades morales no son reducibles a los hechos naturales (interpretándolos por ejemplo, como supervinientes a los hechos naturales).

Por su parte, el intuicionismo social enfrenta dificultades que se siguen de la separación entre los procesos que dependen del sistema intuitivo y los que son producto del sistema racional. Si los juicios morales son la única posición moral que somos capaces de producir, no hay otra alternativa que hacer juicios morales basados en los prejuicios. Haidt proporciona algunas pistas sobre cómo evadir este problema. Dado que el módulo moral es diferente del módulo racional, la manera de modificar las intuiciones morales es suscitar nuevas intuiciones. De modo que sería posible asumir una posición diferente al representarse a uno mismo en otras condiciones sociales o económicas. Los experimentos mentales son herramientas que posibilitarían estas modifi-



caciones porque permiten «gatillar» nuevas intuiciones al proponer escenarios en los cuales el experimentador (o lector del experimento) debe responder a la pregunta «¿cómo debería actuar?» dadas ciertas condiciones antecedentes.

Aunque se propone como una perspectiva que intenta superar las limitaciones del intuicionismo y del intuicionismo social, el racionalismo también enfrenta objeciones que podrían lesionar gravemente su plausibilidad. Por un lado, podría atribuirse al racionalismo la suposición de que la deliberación moral conduce a juicios universales. Esta suposición es un problema en sí misma porque el desacuerdo sobre cuestiones morales es un hecho, pero además involucra una serie de supuestos sobre el sujeto del juicio moral para garantizar tal universalidad; se los supone sujetos máximamente racionales, capaces de evaluar consistentemente los medios para determinados fines, se los considera agentes frónéticos. En respuesta, la perspectiva racionalista propone que, al postular una continuidad entre intuiciones morales y razonamientos, no se apela a un ideal abstracto de razonamiento moral de acuerdo con el cual el razonamiento descansa en recursos e información ilimitada:

De hecho, las operaciones lógicas e inferenciales de conceptos sobre las que estamos hablando se basan solo en piezas específicas de información que una persona ha adquirido sobre el mundo y que es capaz de incluir en una ocurrencia específica de procesamiento de esa información que es en ocasiones muy limitada, inadecuada y recíprocamente, incoherente y, más importante, sobre simplificada y afectada por prejuicios. (Dellantonio y Job, 2012:245)

En suma, las distintas interpretaciones enfrentan desafíos que comprometen su aceptabilidad como buenas explicaciones del rol de las intuiciones en el razonamiento moral. Un punto en común en estas dificultades es explicar el desacuerdo, las intuiciones en conflicto y el cambio de posición moral. En el siguiente apartado examinaremos estos puntos con más detalle.

## **Desacuerdo moral y disidencia**

Uno de los principales obstáculos para las perspectivas intuicionistas surge del hecho del desacuerdo moral. Las variaciones en los juicios morales pueden ser indicativas de la existencia de intuiciones morales en conflicto. Asimismo, una crítica común al intuicionismo moral es su falta de recursos de orden superior para ordenar principios en conflicto (Lariguét, 2017:135). Por otra parte, investigaciones en psicología cognitiva han mostrado que los juicios basados en intuiciones son en muchas ocasiones erróneos, por estar di-

rigidos por sesgos, basarse en información escasa y no ser susceptibles de ser justificados racionalmente. Estas evidencias ponen en cuestión el supuesto de acuerdo con el cual los hechos morales son los hacedores de verdad de los juicios morales. Aunque estas constituyen dificultades innegables para el intuicionismo, existen respuestas plausibles a cada una. Por un lado, algunas perspectivas intuicionistas son compatibles con el pluralismo, de acuerdo con el cual existe una multiplicidad irreductible de hechos morales. Por otro lado, las intuiciones pueden ser reinterpretadas como pareceres intelectuales, por lo que pueden considerarse falibles, es decir, es posible que los sujetos capten incorrectamente principios o propiedades morales. Las interpretaciones aprioristas respecto de las intuiciones en matemática adoptan estrategias similares para evadir esta segunda objeción. Brown por ejemplo,

afirma que, a pesar de que la captación de entidades abstractas es a priori, es posible tener creencias equivocadas acerca del concepto correcto (Brown, 2008:24). Para ilustrar esta afirmación toma un ejemplo de la teoría de conjuntos. Del axioma de comprensión de Cantor, de acuerdo con el cual para cualquier condición hay un conjunto de objetos que la satisface, es posible derivar contradicciones si se intenta construir conjuntos bajo determinadas condiciones, como mostró Russell. La paradoja de Russell condujo, según Brown, a la revisión del axioma y del concepto de conjunto y a un consecuente cambio conceptual.

El intuicionismo social ha sido criticado fundamentalmente por sus limitaciones para explicar la capacidad individual de evaluar de manera autónoma una situación moral. Para esta perspectiva cuando expresamos una posición moral lo hacemos basados en intuiciones, por lo que es difícil dar cuenta de cómo los sujetos son capaces de formular juicios disidentes de aquellos determinados por el marco ideológico del grupo. En respuesta esta consideración, el intuicionismo social propone que, aunque el módulo moral es rígido y encapsulado, es posible que los sujetos modifiquen sus posiciones morales a partir de la persuasión moral. Desde esta perspectiva, la modificación en las posiciones morales no es originada por argumentos lógicamente convincentes, sino por la generación de nuevas intuiciones morales (Haidt, 2001: 819). Este proceso puede ser suscitado por el mero hecho de que amigos, aliados y conocidos hayan emitido un juicio moral (la persuasión social ejerce una influencia directa sobre los sujetos y genera intuiciones morales), a partir de la comprensión empática causada por ponerse en el lugar de otro o debido a argumentos que apelan a disparadores de nuevas intuiciones como los experimentos mentales.

El enfoque racionalista, por su parte, pone su atención el fenómeno de la disidencia que considera como emblemático del razonamiento moral. De acuerdo con este punto de vista, las formas verdaderas de la moral son aquellas capaces de romper con las normas y costumbres de un grupo en orden a seguir diferentes principios. De modo que la disidencia es la condición moral por excelencia: una forma de desacuerdo expresado en cierto punto por un miembro del grupo sobre algún principio o posición perteneciente al marco ideológico de dicho grupo (Dellantonio y Job, 2012:248). Respecto de desacuerdo moral en general, las posiciones intuicionistas son incapaces, desde el punto de vista racionalista, de dar una explicación completamente satisfactoria de este fenómeno. Si bien el intuicionismo social puede argüir que, dado que los juicios morales son dirigidos por principios sociales, las diferencias entre intuiciones morales pueden ser adscriptas a diferencias culturales, debe clarificar cómo y por qué esto pasa: en qué consisten estas diferencias en exposición y asimilación y qué tipo de información entre estas variedades es relevante. Por su parte, la suposición de acuerdo con la cual el módulo moral es irreflexivo e informacionalmente encapsulado, hace difícil de explicar cómo puede reorganizarse y ser sensible a nueva información lo que resulta a su vez fundamental para la explicación del fenómeno de la disidencia.

El principal problema de las posiciones intuicionistas, en sus dos versiones, reside en que, si los juicios morales son producidos por el módulo o el sentido moral que es independiente del sistema de razonamiento, entonces es difícil explicar cómo las personas son capaces de desafiar las intuiciones morales sobre la base de razonamientos morales. Las perspectivas racionalistas proponen que la solución a este problema se encuentra en la tesis continuista, de acuerdo con la cual intuiciones y razonamientos dependen del mismo tipo de procesamiento de la información. En estas formas de razonamiento la información semánticamente estructurada de piezas de información que componen los conceptos puede ser conectada para formar cadenas de inferencias deductivas o inductivas y otras formas más complejas de inferencia como los razonamientos abductivos. Este procesamiento es percibido como instantáneo en el caso de las intuiciones, pero en ocasiones encuentra un *impasse* que puede ser causado por confrontaciones con otras personas o por la necesidad de encontrar un argumento explícito, y en estos casos el razonamiento es explícito, reflexivo y consciente. Los estereotipos pueden ser identificados como principios o intuiciones morales. Entendidos como una forma de conocimiento estructurado, empleado para categorizar grupos o miembros específicos de un grupo, tienden a funcionar basados en un número limitado de atributos, en descuidar diferencias individuales, en no garantizar generalizaciones sobre individuos y grupos. En este sentido los estereotipos

simplifican el procesamiento de información reduciendo la variabilidad en el *input* y recuperando estructuras de conocimiento disponibles. Pero si bien pudieron haber jugado un rol adaptativo, no pueden considerarse desde el punto de vista racionalista como instancias morales positivas, ya que las conclusiones son el resultado de aplicación de prejuicios y en la mayoría de los casos no resisten un análisis posterior. La posición racionalista argumenta que las intuiciones morales espontáneas que personas completamente enculturadas pueden experimentar, por lo general no expresan el mejor juicio posible. Los juicios morales están mejor garantizados cuando no responden intuitivamente a la situación, sino que interrumpen este procesamiento automático y toman el control «manual» en un proceso racional (Dellantonio y Job, 2012:239–240).

Otras perspectivas han sugerido que una manera de resolver los problemas del intuicionismo es proporcionar criterios que permitan resolver conflictos generados por el desacuerdo y jerarquizar principios morales. Rawls ([1971]2006) propone el método del equilibrio reflexivo como manera de resolver desacuerdos y dos principios de justicia, derivados del experimento metal de la situación originaria (Rawls, [1971]2006:135–136) para jerarquizar principios. El primer criterio reza: «cuando los sujetos deliberan correctamente, los juicios hipotéticos acerca de una situación particular *coinciden necesariamente* con los que los sujetos emiten cuando se enfrentan de hecho a la misma» (Rawls, [1971]2006:38). De acuerdo con los segundos: cada persona ha de tener un derecho igual al más extenso sistema total de libertades básicas compatible con un sistema similar de libertad para todos, y las desigualdades económicas y sociales han de ser estructuradas de manera que se garantice la igualdad de oportunidades y que la distribución de las riquezas sea regulada requiriendo que las desigualdades económicas sean aquellas que redunden en beneficio de todos, en particular, de los menos aventajados. Estas estrategias han sido criticadas principalmente sobre la base de que el equilibrio reflexivo solo puede garantizarse asumiendo una serie de suposiciones respecto de los sujetos que juzgan que son determinantes del resultado del proceso deliberativo, fundamentalmente: que los sujetos pueden hacer abstracción de las condiciones a partir de las cuales deliberan (rasgos biológicos, posición social, determinaciones ideológicas particulares), que todos los sujetos han de gozar de los mismos derechos a la hora de participar en los procedimientos mediante los que se determinan los principios de la justicia y que poseen ciertas virtudes epistémicas, entre ellas, la capacidad de determinar los medios más efectivos para un fin y la capacidad para una concepción del bien independiente de aquella con la que cada uno se identifica. Una crítica interesante a los supuestos que sostienen la posición rawlsiana puede encontrarse en el trabajo

de Corbí (2010) donde se muestra que existen importantes divergencias entre los juicios que pueden alcanzarse como producto de un proceso deliberativo en abstracto, y aquellos que son producto de procesos de razonamiento que involucran los sujetos que estén inmersos en ciertas experiencias. De acuerdo con Corbí, solo los segundos pueden considerarse juicios maduros y meditados, los primeros son el resultado de juicios apresurados y confusos.

En síntesis, el razonamiento moral desempeña, desde algunas perspectivas, un rol importante en los casos de desacuerdo y la disidencia explicando cómo es posible que los sujetos cambien su punto de vista respecto de un asunto y posibilitando una vía para el ajuste de posiciones en conflicto. Por el contrario, las intuiciones, como respuestas automáticas, han sido asociadas a prejuicios o estereotipos. Se arguye que presentan información fragmentada, parcial e incompleta y por lo tanto que su papel en los juicios morales es negativo. No obstante, es posible considerar casos en los que las intuiciones juegan un rol positivo en la transformación de posiciones morales y favorecen el consenso cuando se presentan desacuerdos. Elaboraremos brevemente esta

## **Optimismo**

Tomando en cuenta que, como sugiere Haidt (2001), la comprensión empática de una situación puede suscitar nuevas intuiciones morales y que, como señala Corbí (2010), las deliberaciones morales en ocasiones presentan importantes diferencias con los juicios emitidos por los sujetos inmersos en ciertas experiencias,<sup>8</sup> proponemos a continuación dos maneras en las cuales las intuiciones, ya sea que estas se entiendan en continuidad con el razonamiento o como una facultad independiente de ella, juegan un papel positivo en los juicios morales facilitando la aceptación de un juicio moral o permitiendo evaluar principios morales.

Por un lado, las intuiciones producto de procesos de enculturación no siempre remiten a prejuicios, sino que en ocasiones son el resultado de luchas sociales, actos de disidencia o razonamientos en sentido amplio (colectivamente generados, producto de la incorporación del razonamiento sobre dilemas en la educación). Puede argüirse que el surgimiento de nuevas intui-

---

<sup>8</sup> Corbí refiere la diferencia entre el juicio del soldado que ha experimentado los horrores de la guerra, como el juicio de un sujeto inmerso en la experiencia. Sugiere que este es sustancialmente diferente de aquel que pudo haber pronunciado el mismo sujeto antes de tener tales experiencias o de los juicios morales respecto de la responsabilidad moral del soldado que otros sujetos, que no experimentaron los horrores de la guerra. Conceptualiza esta diferencia como la diferencia entre la conciencia declarativa y la conciencia expresiva (Corbí, 2010:98).

ciones morales que juegan un rol en los juicios y en el ajuste o revisión de las creencias en los sujetos que las tienen son un ejemplo del papel positivo de las intuiciones morales. Largos procesos de vindicación de derechos como los llevados a cabo por los movimientos feministas han suscitado intuiciones que tienen un rol en la conducta moral, por ejemplo, inhibiendo ciertos comportamientos (como comentarios sobre el aspecto físico de las mujeres, actitudes de ridiculización o menosprecio) o propiciando actitudes empáticas (que dependen del reconocimiento de diferencias materiales y simbólicas). En estos casos puede considerarse que nuevas intuiciones surgidas de las luchas sociales que han expuesto la opresión y exclusión sistemática de las mujeres y las múltiples formas de violencia asociadas al machismo, propician formas de conducta basadas en juicios automáticos que censuran prácticas discriminatorias y tratos injustos. Aunque no puede decirse que se trata de intuiciones universales, sino que su incorporación es progresiva y está determinada por la influencia que dichas luchas sociales han tenido en cada contexto local. Algo similar puede afirmarse respecto de intuiciones morales relativas al sufrimiento animal que motivan actitudes compasivas o incluso conductas de cuidado hacia los animales. En general, las discusiones instaladas en la agenda pública por los movimientos ecologistas que de un tiempo a esta parte se han incorporado también en la educación de las generaciones jóvenes, han provocado intuiciones morales que determinan conductas de respeto con el medio ambiente y propician la percepción de ciertas conductas como deberes morales.

Otro modo en que las intuiciones pueden influir positivamente en los razonamientos morales está representado por su rol en los experimentos mentales. Una tesis, ampliamente aceptada, es que los experimentos mentales son una estrategia metodológica que funciona sobre la base de intuiciones. Los escenarios presentados en la narrativa de un experimento mental suscitan intuiciones (entendidas estas como inclinaciones a creer) que hacen plausible o que socavan la credibilidad de una hipótesis. En el análisis de estos casos, las intuiciones pueden favorecer una reestructuración en la jerarquía de los principios morales o posibilitar la revisión de principios hasta el momento no cuestionados. Los procedimientos de suposición empleados en los experimentos mentales pueden favorecer además procesos de consenso en casos de desacuerdo entre pares epistémicos. Los escenarios hipotéticos posibilitan la formulación de un problema desde una perspectiva desde la cual no había sido presentado. En algunos casos, el escenario hace posible «desactivar» una posición idiosincrática que obtura algunas distinciones conceptuales importantes. Tomemos como ejemplo el experimento mental que Judith Thomson (1971:48-49) propone para discutir la permisibilidad moral del aborto.

El caso plantea un escenario en el cual un individuo se ve sorpresivamente conectado a otro y que debe permanecer en dicha situación durante nueve meses para mantener viva a la otra persona. El experimento estipula un escenario en el cual es posible separar, al menos teóricamente, la obligación moral de salvar una vida, del derecho de disponer aquello necesario para sustentarla. El caso tiene la virtud de poner al lector en la posición de protagonista. El planteo sugiere que al encontrarse en una situación como la descrita el lector podría optar por desconectarse del violinista sin avasallar con eso los derechos de este último. Aquello que intuitivamente parece justo hacer en el caso, debería valer (bajo el supuesto de que el excepcionalismo de primera persona está moralmente prohibido) para otros casos.<sup>9</sup>

## Conclusiones

La naturaleza de las intuiciones continúa siendo materia de controversia en la filosofía y en las disciplinas científicas vinculadas a la ciencia cognitiva. El trabajo colaborativo entre científicos y filósofos ha iluminado uno de los problemas filosóficos más acuciantes: la relación mente/cuerpo. Sin embargo, no es posible afirmar que estos avances hayan permitido responder de manera definitiva las preguntas que motivan el análisis filosófico de las intuiciones. Una manera de presentar la cuestión que consideramos motiva este examen es la siguiente: ¿son las intuiciones vías de acceso a verdades, objetos o propiedades o son meros espejismos de nuestros prejuicios? Aunque en el ámbito

---

<sup>9</sup> La descripción del contrafáctico sugiere una interpretación teórica o la vuelve más factible porque «gatilla» ciertas intuiciones e «inhibe» otras. Las posiciones optimistas respecto de los poderes epistémicos de los experimentos mentales son capaces de responder parcialmente al reto que plantean las perspectivas escépticas asumiendo que, si bien es posible que estos dispositivos propicien respuestas inadecuadas, pueden ser herramientas metodológicas útiles si se ciñen a criterios de fiabilidad que restrinjan la proliferación de representaciones sugestivas o irrelevantes para evaluar una hipótesis. Aunque no existe un consenso generalizado respecto de estos criterios, se pueden identificar dos en relación con las objeciones formuladas por las posiciones escépticas. En general, el uso legítimo de los experimentos imaginarios en el marco de una controversia o disputa filosófica debería estar sujeto a la condición de que empleo solo es legítimo en los casos en los que el punto de vista del oponente es establecido con claridad y cualquier idealización del argumento defendido sea concedido o aceptado por el oponente (esta idea es sugerida por Popper, 1959). Esto garantiza que el experimento no caricature ni simplifique excesivamente las posiciones en discusión. En segundo lugar, el escenario imaginario empleado en el experimento no debe estar subdeterminado. Esto implica que la manipulación (abstracción, eliminación) de variables debe estar bajo control, es decir, debe ser claro cómo se vinculan con las suposiciones involucradas y cómo se relacionan (apoyan) con la conclusión (Brendel, 2005). Aunque estos criterios son un tanto laxos, proporcionan una base de calibración para identificar usos espurios de los experimentos mentales.

del conocimiento matemático y físico las intuiciones parecen estar asociadas a destellos de claridad, defender que estas constituyen una forma de conocimiento alternativa a la percepción y el razonamiento conlleva un costo metafísico difícil de solventar.

En el contexto de la discusión ética y metaética, el examen de las intuiciones despliega otras dimensiones. Elucidar la naturaleza y funcionamiento de las intuiciones en el razonamiento moral haría posible elucubrar modos de resolver el desacuerdo moral, suspender el juicio o limitar las controversias a las cuestiones que sean susceptibles de resolución. Entre las cuestiones abiertas que consideramos deben ser respondidas para progresar en este sentido identificamos tres: el problema de la justificación de las intuiciones, la relación entre intuiciones y estereotipos, y el vínculo entre intuiciones morales y emociones. El primer problema concierne al rol y estatus de las intuiciones en la cadena de justificación. Si las intuiciones pueden proporcionar garantías para juicios normativos es necesario establecer claramente de qué manera lo hacen. La segunda cuestión es especialmente atinente a las intuiciones morales. Consiste en determinar en qué medida las intuiciones son reducibles a prejuicios o estereotipos como respuestas instantáneas. En este punto, nuestra contribución consiste en plantear casos en los cuales las intuiciones, como juicios abreviados o respuestas automáticas, podrían cumplir un rol positivo no solamente traccionando el razonamiento hacia principios socialmente incorporados, sino como mecanismo que puede ampliar las consideraciones del agente, por ejemplo, a través de la empatía. El tercer problema abierto que estimamos demanda atención para realizar progresos en las discusiones sobre las intuiciones morales, es el de la relación entre emociones e intuiciones. Si bien en este trabajo no abordamos directamente este punto, la perspectiva del intuicionismo social que encuentra sus bases en el emotivismo humeano, señala que las intuiciones son experimentadas como sentimientos viscerales; consideramos que cabe preguntarse si es posible separar las emociones de las intuiciones o de las reacciones morales en general, si es posible identificar procesos de razonamiento moral que son puramente cognitivos y cuál es el vínculo entre intuiciones, emociones y razonamiento.



## Referencias bibliográficas

- Alexander, J. y Weinberg, J. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass*, 2(1), 56–80.
- Allman, J. y Woodward, J. (2008). What Are Moral Intuitions and Why Should we Care about Them? A Neurobiological Perspective. *Philosophical Issues*, 18, 164–185.
- Audi, R. (2013). *Moral perception: Soochow lectures in philosophy*. Princeton University Press.
- Bengson, J. (2015). The Intellectual Given. *Mind*, 124(495), 707–760.
- Boghossian, P. (2014). Philosophy without Intuitions? a Reply to Cappelen. *Analytic Philosophy*, 55(4), 368–381.
- Brendel, E. (2004). Intuition Pumps and the Proper Use of Thought Experiments. *Dialectica*, 58(1), 89–108.
- Brown, J. (2011). *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. Second Edition. Routledge.
- Brown, J. (2008). *Philosophy of Mathematics*. Second Edition. Routledge.
- Buzzoni, M. (2009). Empirical Thought Experiments: A Transcendental–Operational View. *Epistemología*, 33(1), 5–26.
- Corbí, J. (2010). El refugio de la claridad. *Análisis Filosófico*, 30(1), 89–121.
- Dellantonio, S. y Job, R. (2012). Moral Intuitions vs. Moral Reasoning. A Philosophical Analysis of the Explanatory Models Intuitionism Relies On. En Magnani, L. y Li, P. (Eds.), *Philosophy and Cognitive Science* (pp. 239–362). Springer.
- Descartes, R. ([1641]2011). Meditaciones metafísicas seguidas de objeciones y respuestas. En *Descartes* (trad. de Jorge Aurelio Díaz) (pp. 153–414). Gredos.
- Descartes, R. ([1628]2011). Reglas para la dirección del espíritu. En *Descartes* (trad. de Luis Villoro) (pp. 3–71). Gredos.
- Enoch, D. (2011). *Taking morality seriously: A defence of moral robust realism*. Oxford University Press.
- Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108(4), 814–834.
- Hume, D. ([1748]2010). *Investigación sobre el conocimiento humano* (trad. de Jaime de Salas Ortueda). Alianza.
- Kalpokas, D. (2017). Percepción y Realismo moral. *Ética y Discurso*, 2(2), 31–44.
- Kant, I. ([1781]2007). *Crítica de la Razón pura* (trad. de Mario Caimi). Colihue.
- Kauppinen, A. (2018). Moral Sentimentalism. En Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*.
- <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=moral-sentimentalism>
- Lariguet, G. (2017). Intuicionismo y razonamiento moral. *Revista de la facultad de derecho*, (79), 127–150.
- Linnebo, Ø. (2017). Platonism in the Philosophy of Mathematics. En Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/platonism-mathematics/>
- Mach, E. ([1897]1976). On Thought Experiments. En Mach, E., *Knowledge and Error* (pp. 134–147). Reidel.

- McDowell, J. (1998). Virtue and Reason. En *Mind, value and reality* (pp. 50–74). Harvard University Press.
- Popper, K. ([1959]2002). On the Use and Misuse of Imaginary Experiments, Especially in Quantum Theory. En Popper, K., *The Logic of Scientific Discovery* (pp. 464–480). Routledge.
- Price, R. ([1758]1969). *A Review of the Principle Questions in Morals*. En Raphael, D.D. (Ed.), *The British Moralists 1650–1800, II* (pp. 131–198). Clarendon Press.
- Rawls, J. ([1971]2006). *Teoría de la Justicia* (trad. de María Dolores Gonzáles). Fondo de Cultura Económica.
- Ross, D. (2001). *Lo correcto y lo bueno* (trad. de L. Rodríguez Duplá, 2da ed.). Sígueme Salamanca.
- Rowbottom, D. (2014). Intuitions in Science: Thought Experiments as Argument Pumps. En Booth, A. y Rowbottom, D. (Eds.), *Intuitions* (pp. 119–134). Oxford University Press.
- Shafer-Landau, R. (2005). *Moral realism: A defence*. Oxford University Press.
- Sosa, E. (2007). Intuitions: Their Nature and Epistemic Efficacy. *Grazer Philosophische Studien*, (64), 51–67.
- Stich, S. y Tobia, K. (2018). Intuitions and its Critics. En Brown, J.; Stuart, M. y Fehige, Y. (Eds.), *The Routledge Companion to Thought Experiments* (pp. 445–442). Routledge.
- Stratton-Lake, P. (2020). Intuitionism in Ethics. En Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=intuitionism-ethics>
- Tallant, J. (2012). Intuitions in physics. *Synthese*, 190(15), 2959–2980.
- Thomson, J. (1971). A Defense of Abortion. *Philosophy and Public Affairs*, 1(1), 47–66.
- van Roojen, M. (2014). Moral Intuitionism, Experiments, and Skeptical Arguments. En Booth, A. y Rowbottom, D. (Eds.), *Intuitions* (pp. 148–164). Oxford University Press.
- Whithead, C. ([1999]2000). *The intuitionist*. Anchor Books.

## **II. Problemas característicos**

## 9. Percepción moral: algunas clarificaciones y distinciones

*Daniel E. Kalpokas*

Se reconoce, en general, que la experiencia perceptiva desempeña un papel esencial en la adquisición y justificación del conocimiento empírico. En tal sentido, se dice, por ejemplo, que las experiencias perceptivas justifican, o constituyen razones para sostener, juicios y creencias empíricas. Ahora bien, suponiendo que haya algo así como «el conocimiento del mundo moral», cabe preguntar: ¿tiene la percepción un rol epistemológico semejante en la justificación de nuestros juicios y creencias morales? Esto es, ¿puede sostenerse que percibimos hechos morales y que esta percepción provee de justificación a nuestros juicios y creencias morales? Recientemente, varios son los filósofos que han argumentado en favor de una respuesta afirmativa a esta pregunta (Bloom, 1994; Greco, 2000; Audi, 2010, 2013, 2015; Dancy, 2010; Chappell, 2008; Cowan, 2015; Cullison, 2009; McBrayer, 2010; McGrath, 2004; Werner, 2016; Wisnewski, 2013).<sup>1</sup> La tesis de que también en el ámbito del conocimiento moral la experiencia perceptiva tiene un rol epistémico que desempeñar se enmarca dentro de un renovado interés por el intuicionismo ético (Huemer, 2008; Stratton-Lake, 2003; Audi, 2004). De ahí que la posición que sostiene que podemos percibir algunos, al menos, hechos morales sea llamada, en ocasiones, «intuicionismo perceptivo» (Cowan, 2015).

El realismo moral está inextricablemente implicado por la tesis de que, al menos en ocasiones, percibimos ciertos hechos morales, pues, si es verdad que percibimos hechos morales, entonces también debe serlo que tales hechos existen: no se puede percibir, en el sentido factivo de «percibir», lo que no existe. En efecto, si, como muchos filósofos están dispuestos a sostener, los verbos de percepción (ver, oír, etc.) son factivos (o de éxito),<sup>2</sup> entonces, de «Percibo *O*» (un hecho moral cualquiera) se sigue que *O* existe. Sin embargo, la conversa no es verdadera: del hecho de que el realismo moral sea verdadero no podemos inferir que la tesis de la percepción moral también lo sea. Podría ser que nuestro conocimiento de los hechos morales fuera enteramente

---

1 Con frecuencia se suele citar, como primer antecedente, el siguiente pasaje de la *Ética Nicomaquea* de Aristóteles: «La prudencia se refiere al otro extremo, a lo más particular, de lo que no hay ciencia, sino percepción sensible, no la percepción de las propiedades sensibles, sino una semejante a aquella por la que vemos que la última figura en matemática es un triángulo», Aristóteles, *EN*: 1142a, 27–29. Para una interpretación de este aspecto de la filosofía de Aristóteles, puede verse Woods (1986).

2 Véase Ryle (1949), Austin (1962), Sellars (1997), Williamson (2000).

de naturaleza inferencial, o que tales hechos fueran conocidos por algún otro medio (v.g., intuición intelectual, etc.). Como muchos otros, en lo que sigue supondré que hay efectivamente hechos morales tales como acciones buenas, valientes, justas o cobardes; temperamentos o rasgos de carácter que son buenos, malos, etc., e instituciones que pueden ser consideradas desde un punto de vista moral. Con todo, en este artículo me concentraré exclusivamente en el caso de las acciones morales.

## 1.

La tesis de la percepción moral podría parecer *prima facie* implausible. ¿Cómo es que podríamos llegar a percibir, por ejemplo, que cierto acto es cruel, valiente o injusto o, más en general, que cierta acción es incorrecta o mala? Parece evidente que no percibimos tales propiedades morales (v.g., la crueldad, la valentía, la injusticia, la incorrección o la maldad) tal como podemos percibir colores, formas, u objetos tales como árboles, tomates o sillas. Y, ciertamente, no han faltado quienes han expresado su escepticismo al respecto (Hume, [1738]1984; Harman, 1977; Mackie, 1977). Los filósofos que defienden la tesis de la percepción moral suelen apelar a distintas estrategias para argumentar en favor de ella. Por ejemplo, algunos apelan, explícita o implícitamente, a ciertos argumentos por analogía. (Audi, 2010, 2013, 2015; Cullison, 2009). En este sentido, suelen decir que, así como parece plausible sostener que, al menos en ciertas ocasiones y para el caso de ciertos estados mentales, podemos percibir directamente los estados mentales ajenos (v.g., cuándo una persona está enojada o siente dolor), así también resulta plausible sostener que podemos percibir que una persona ha hecho algo malo. En otros casos, no obstante, se argumenta indirectamente en favor de la tesis de la percepción moral mostrando que las tesis rivales (v.g., la tesis de que conocemos los hechos morales mediante inferencia a la mejor explicación) *no* pueden abarcar *todos* los casos de conocimiento moral. Al menos en algunas circunstancias, tenemos que reconocer que ciertos hechos morales son conocidos perceptivamente (McGrath, 2004). Más recientemente, se han avanzado argumentos directos en favor de la tesis de la percepción moral, basados en la fenomenología de la experiencia (Werner, 2016). En este trabajo, quisiera hacer explícitos los compromisos teóricos que involucra la tesis de la percepción moral. Para ello, introduciré algunas distinciones que no siempre se tienen en cuenta en la literatura, y algunas clarificaciones que conciernen al contenido de la percepción moral y al modo en que se supone que esta (re) presenta los hechos morales. El objetivo último de todo esto es presentar del modo más claro, explícito y plausible posible dicha tesis.

## 2.

Comencemos con algunas distinciones preliminares que es preciso considerar, antes que nada. ¿Qué se quiere significar exactamente al decir que, al menos en ciertas ocasiones, podemos tener experiencia perceptiva de hechos morales? El término «experiencia» posee, ciertamente, distintos sentidos, o se usa comúnmente de diferentes maneras. Por ejemplo, tener dolor de muelas es un tipo de experiencia. Sentir indignación ante, por ejemplo, una injusticia, es otro tipo de experiencia. También se habla de tener experiencia en, pongamos por caso, un trabajo o una profesión. Ninguno de estos usos es relevante para especificar la tesis de la percepción moral.<sup>3</sup> Más bien, la idea es que, según dicha tesis, tenemos, o podemos tener, experiencia sensorial o perceptiva de ciertos hechos morales, podemos percibir, literalmente hablando, que ciertas acciones son valientes, cobardes, crueles o, más en general, correctas, buenas o malas.<sup>4</sup>

Conviene hacer dos aclaraciones respecto de la percepción moral así entendida. Algunos filósofos —especialmente, McDowell (1998)— han hablado de percepción moral como aquella capacidad que los sujetos virtuosos tienen de percibir algunos rasgos salientes de las situaciones particulares en las que se encuentran, rasgos que, en tanto percibidos, ayudan a explicar por qué actúan como lo hacen en tales circunstancias. Este es, ciertamente, un sentido de «percepción» que se parece bastante al que puede hallarse en la obra aristotélica. Sin embargo, este no parece ser exactamente el sentido en el que los filósofos de hoy en día hablan de «percepción moral». La percepción moral en el sentido antedicho parece ser la capacidad que los sujetos virtuosos poseen de detectar o advertir perceptivamente qué acción ha de ser llevada a cabo, en circunstancias determinadas, para alcanzar cierto bien. La tesis de la percepción moral, en el sentido en el que nos interesa aquí, remite, en cambio, a la capacidad de identificar perceptivamente ciertos hechos morales como, por ejemplo, una acción justa, valiente, o incorrecta moralmente hablando. Seguramente, ambos sentidos están estrechamente relacionados, pero, mientras que uno de ellos enfatiza, desde la perspectiva del agente, la percepción de ciertas oportunidades (o *affordances*) para hacer lo que es correcto en determinadas circunstancias, el otro, en cambio, enfatiza, desde la perspectiva de un sujeto cognoscente, el conocimiento moral no inferencial de ciertos hechos que puede adquirirse por el simple hecho de percibirlos.

---

3 Con todo, no quisiera negar la posibilidad de que la percepción moral pueda ir acompañada de ciertos sentimientos o emociones, ni que esto pueda también afectar la fenomenología de la experiencia moral.

4 Para un estudio sobre los distintos usos del término «experiencia», puede consultarse Hinton (1973).

Naturalmente, el defensor de la tesis de la percepción moral ha de tomar una serie de decisiones teóricas acerca de la naturaleza de la experiencia perceptiva. Para comenzar, ha de tomar partido en la disputa entre los que sostienen que la experiencia perceptiva posee contenido representacional o intencional (McDowell, 1996; Searle, 1983; Tye, 1995; Crane, 1992; Siegel, 2010) y los que sostienen todo lo contrario: esto es, que no posee contenido representacional alguno. Para estos últimos, la percepción es, *meramente*, una relación consciente, inmediata, entre el sujeto y el objeto percibido (Campbell, 2002; Travis, 2004; Brewer, 2006, 2008, 2011; Martin, 2002, 2009; Johnston, 2006; Fish, 2009; Raleigh, 2013). Aquí, en mi opinión, es difícil ver cómo podría sostenerse alguna versión de la tesis de la percepción moral si prescindiéramos de la idea de que la percepción es representacional. Lo que estamos buscando, en realidad, es hacer plausible la idea de que al menos algunos contenidos representacionales, de carácter específicamente moral, figuran en el contenido de nuestros estados perceptivos. De otro modo, no se entendería cómo la percepción podría hacer algún aporte epistémico al conocimiento de los hechos morales. Quizá precisamente por esto es que todos los intentos de defender la tesis de la percepción moral han supuesto, hasta ahora, que la experiencia perceptiva es representacional. Obviamente, no puedo embarcarme aquí en una discusión acerca del carácter representacional o no representacional de la experiencia. Habiendo destacado esta distinción, asumiré directamente en este punto que la percepción tiene, efectivamente, contenido representacional.

Por otro lado, una vez reconocido el carácter representacional de la experiencia, es preciso tener en cuenta el debate en torno a la naturaleza conceptual o no conceptual del contenido perceptivo.<sup>5</sup> También aquí son muchos los que directamente suponen que, para representarse perceptivamente ciertos hechos morales, el sujeto ha de contar con ciertas capacidades conceptuales que han de figurar en el contenido de su experiencia. Nuevamente, resulta difícil entender aquí cómo podría alguien ver que cierto acto es injusto, por ejemplo, o bueno, si en el contenido de su experiencia no figuraran los conceptos mismos de injusticia o de bondad. De cualquier modo, los mismos defensores del no-conceptualismo han restringido su tesis para el caso de la percepción de colores, formas, etc.; esto es, no han osado llevarla tan lejos como para afirmar que también para el caso de la percepción de libros, bancos o acciones morales el contenido de la experiencia puede ser de carácter no

---

5 Sobre este debate, véase, por ejemplo, McDowell (1996), Brewer (1999), Peacocke (1992), Evans (1982), Dretske (1981), Cussins (2003), Heck (2000).

conceptual (Peacocke, 1992). Otra vez, pues, seguiré a la mayoría al sostener que el contenido perceptivo es de naturaleza conceptual.<sup>6</sup>

Otra disputa que también es relevante para la presente discusión es la que existe en filosofía de la percepción acerca de cuáles son las propiedades que, efectivamente, son representadas perceptivamente. La discusión gira en torno a la distinción entre propiedades de bajo nivel (v.g., color, forma, etc.) y propiedades de alto nivel (v.g., clases de objetos, causalidad, etc.). Aquí, el defensor de la tesis de la percepción moral no tiene muchas alternativas, pues difícilmente pueda sostenerse que las propiedades morales (v.g., justicia, bondad, etc.) son de bajo nivel. Así pues, si la tesis de la percepción moral ha de tener plausibilidad, ha de adoptarse la tesis de que también las propiedades de alto nivel pueden ser representadas en la percepción.<sup>7</sup>

Finalmente, parece que también es importante tener presente la idea de percepción a-modal. Aunque no es del todo claro el alcance ni la naturaleza de este fenómeno, podemos caracterizar la percepción a-modal siguiendo a Nanay: «Percibimos a-modalmente una parte de un objeto (percibido) si no recibimos estimulación sensorial de esa parte del objeto» (Nanay, 2010:241-242).<sup>8</sup> La idea es que, al percibir visualmente un objeto sólido, en cierto sentido solo vemos, en cada momento, cierta parte de él (v.g., su cara de enfrente). No obstante, nos parece como si, fenoménicamente hablando, la cara visible tuviera una cara posterior. Es por esto que, al ver la cara visible de un tomate, digamos, nos parece que estamos viendo un objeto tridimensional. La parte de atrás del tomate, que en cierto sentido no podemos ver mientras vemos la parte de adelante, es representada a-modalmente, según se dice, y, así, junto con la representación de la cara visible, esta representación de la cara posterior también forma parte del contenido total de la experiencia.<sup>9</sup> Ahora bien, a fin de hacer plausible la tesis de la percepción moral, algunos filósofos han apelado a la percepción a-modal (Cowan, 2015).<sup>10</sup> La idea es que la representación perceptiva a-modal podría explicar, quizá, la diferencia intuitiva que parece haber entre la percepción de colores y formas, por un lado, y la percepción de propiedades morales, por otro (volveré sobre este punto en § 6).

---

6 Argumento en favor del conceptualismo en Kalpokas (2018).

7 Esto, a su vez, conduce a suponer, además, la tesis de la penetrabilidad cognitiva: la tesis de que ciertos estados cognitivos pueden causar ciertas modificaciones en el contenido y la fenomenología de la experiencia perceptiva. Para la tesis según la cual en la experiencia perceptiva representamos propiedades de alto nivel, véase Siegel (2006, 2010), Bayne (2010).

8 Para la tesis de la percepción a-modal, véase también Noë (2004), Nanay (2013), Briscoe (2011).

9 Por cierto, hay distintas explicaciones de este fenómeno. Para mis presentes propósitos, bastará con la caracterización general ofrecida por Nanay. Para distintas explicaciones de la percepción a-modal, véanse los trabajos citados en la nota anterior.

10 En lugar de percepción a-modal, Audi (2013) habla de percepción «intelectiva».



Con respecto a todos estos puntos que he señalado, que conciernen a distintos debates en epistemología y filosofía de la percepción, parece que, a fin de articular una propuesta plausible acerca de la percepción moral, conviene decantarnos por la idea de que la percepción posee contenido representacional, de carácter conceptual, que involucra la representación de propiedades de alto nivel y, quizá, la percepción a-modal de aspectos de la realidad que, al menos en ciertos momentos, no nos resultan directamente perceptibles. Antes de esbozar un argumento que, según mi opinión, es suficiente para darle plausibilidad a la tesis de la percepción moral, es preciso, todavía, aclarar qué tipos de conceptos son los que supuestamente podrían figurar en el contenido de la experiencia moral. En este punto, me refiero a la distinción entre conceptos «densos» (*thick*) y conceptos «finos» o «delgados» (*thin*). La distinción proviene de Williams (1985). Los conceptos densos son aquellos que, en su contenido, combinan elementos descriptivos y evaluativos, como, por ejemplo, *generoso*, *cruel*, *valiente*, o *amable*. Se trata de conceptos que parecen ser descriptivamente más específicos o ricos que conceptos delgados tales como *bueno*, *malo* o *incorrecto*, que parecen ser más generales o puramente evaluativos.<sup>11</sup> Ahora bien, si la versión más plausible de la tesis de la percepción moral ha de involucrar la tesis conceptualista acerca del contenido perceptivo, cabe preguntarse: los conceptos que podrían figurar en el contenido de nuestras supuestas experiencias morales, ¿serían conceptos densos o delgados? ¿O ambos?

### 3.

Permítaseme presentar ahora un argumento en favor de la percepción moral basado en el carácter fenoménico de la experiencia perceptiva que resulta, según pienso, atendible. En la literatura sobre el tema, con el objeto de construir un argumento basado en la fenomenología de la percepción moral, se suele apelar a lo que Siegel (2006, 2010) llama «el método del contraste fenoménico». La idea consiste en contrastar dos experiencias que, aunque refieren al mismo objeto, estado de cosas o escena, poseen diferente carácter fenoménico. El método procede buscando la mejor explicación de esa diferencia fenoménica. Se argumenta que la mejor explicación es aquella que sostiene que la diferencia fenoménica se debe a que solo en un caso la experiencia representa una propiedad de alto nivel como, por ejemplo, una propiedad mo-

---

11. Para una visión actualizada de la discusión en torno a los conceptos densos y delgados, véanse los artículos compilados en Kirchin (2013).

ral. El hecho de que una experiencia represente, en efecto, tal propiedad es lo que explica por qué posee un carácter fenoménico que la distingue de la otra experiencia con la que es contrastada, que, se supone, no representa tal propiedad.<sup>12</sup>

Ahora bien, me parece que este es un método interesante y útil para investigar qué propiedades de alto nivel son efectivamente representadas en la experiencia (cualesquiera sean esas propiedades). Sin embargo, pienso que para hacer plausible la idea de la percepción moral basta con un argumento más modesto y, en cierto sentido, «previo» al del contraste fenoménico. *No* es necesario probar que el caso bajo consideración es, *de hecho*, un caso de percepción moral. Sin duda, podría serlo; pero —y este es el punto— también podría no serlo. Quiero decir, podría tratarse de un caso en el que solo *parece*, fenoménicamente hablando, que estamos percibiendo un hecho moral. Y esto, sostengo, es suficiente para mostrar que algunos hechos morales *podrían* ser representados perceptivamente. El hecho de que pueda parecerle a uno como si estuviera percibiendo un hecho moral basta para contrastarlo con un caso en el que ni siquiera puede parecerle a uno que está percibiendo un hecho moral. Así, mientras que la estrategia del contraste fenoménico intenta explicar las diferencias en la fenomenología de la experiencia moral, el argumento que presentaré a continuación pretende mostrar que esa diferencia es genuinamente fenoménica, esto es, que no puede ser reducida a una diferencia en los juicios, o creencias, que las personas se forman acerca de los hechos morales percibidos.

Consideremos el ejemplo de Harman (1977), que se ha puesto de moda en las discusiones sobre percepción moral. Supongamos que vemos a un grupo de niños rociando con nafta a un gato y prendiéndole fuego. Ni bien vemos la escena, advertimos que lo que los niños están haciendo es incorrecto. Pero, ¿se debe nuestra reacción al hecho de que vemos que lo que los niños están haciendo es incorrecto? La respuesta de Harman es que no, que la incorrección del acto no puede ser observada en absoluto. Sin embargo, si apelamos a la fenomenología de la experiencia perceptiva, cabe argumentar todo lo contrario.

Para empezar, es importante especificar en qué tipo de error moral están incurriendo los niños al prenderle fuego al gato. Se trata a todas luces de un acto de crueldad. Los términos «crueldad» y «cruel» son términos morales densos. La crueldad se define como la propiedad de causar intencionalmente dolor y sufrimiento innecesarios a otros seres vivos. Así, podemos suponer que lo que están haciendo los niños en el ejemplo de Harman instancia dicha propiedad. ¿Podemos observar esta instanciación en sus acciones? ¿Cómo se-

---

12 Werner (2016) apela a esta estrategia.

ría ver una acción cruel? ¿Qué aspecto tendría? La sugerencia consiste en concebir la crueldad (y, en general, otras propiedades morales) como una propiedad que puede instanciarse en un patrón, o estilo, de comportamiento e interacción característico.<sup>13</sup> Vemos a los niños prendiéndole fuego al gato y vemos el gato en llamas, corriendo a los gritos. Reconocemos visualmente que los niños están causando dolor al gato innecesariamente. Justamente, su patrón conductual en referencia al gato satisface las condiciones de aplicación de nuestro concepto de crueldad. Así, podemos decir que identificamos, o reconocemos, que su patrón conductual instancia la propiedad de la crueldad. No inferimos la crueldad de sus acciones a partir de su observación; más bien, vemos directamente que su patrón conductual ejemplifica el patrón que constituye la condición de aplicabilidad de nuestro concepto de crueldad.

Lo interesante del caso es que, aun si no se tratara de una situación real, nuestra experiencia podría tener la misma fenomenología. En efecto, supongamos por un momento que la escena vista no es real en el sentido de que se trata, por ejemplo, de una escena de una película que se está filmando. En verdad, los niños no están torturando a ningún ser vivo; se trata de un truco de filmación. Si esto es así, el patrón conductual de los niños no está instanciando la propiedad de la crueldad, aunque pueda parecernos, visualmente, que lo hace. El punto es que, aunque nuestra experiencia no esté representando verdaderamente la propiedad de la crueldad, es como si, fenoméricamente hablando, lo estuviera haciendo. Estaríamos en un caso de ilusión visual moral: las cosas no son como nos parecen que son. Ahora, si cierto patrón de comportamiento, que no instancia realmente la crueldad, puede parecernos que sí la instancia, es porque debe existir cierta similitud entre ambos. En general, no puede parecerle a uno, engañosamente, que  $x$  instancia la propiedad  $F$  si uno no tiene la menor idea de cómo es que a uno le parezca, verdaderamente, que  $x$  tiene la propiedad  $F$ . Por tanto, si es posible que haya algo así como ilusiones visuales morales, entonces existen algunas bases para sostener que, al menos en algunos casos, la fenomenología de nuestra experiencia nos presenta algunas conductas como instanciando ciertas propiedades morales. Aunque podemos suponer, como en el caso de la experiencia perceptiva empírica, que la mayoría de las veces nuestras experiencias representan efectivamente las propiedades que parecen representar, no es preciso probar que de hecho ese es el caso para el ejemplo considerado. Basta con que nos parezca, fenoméricamente hablando, que nuestra experiencia representa la crueldad como instanciada en las acciones de los niños para que podamos hacer plau-

---

13 Chappell (2008) enfatiza el rol central que desempeñan los patrones conductuales en la percepción moral.

sible la idea de que la experiencia perceptiva puede representar, al menos en ocasiones, acciones crueles.

Para reforzar este punto, podemos suponer ahora que somos advertidos por los productores de la película que lo que estamos viendo, o estamos por ver, solo es una escena dentro de la película. Y podemos imaginar que la escena es tan vívida que, nuevamente, nos parece visualmente como si los niños estuvieran cometiendo un acto de inmensa crueldad con el gato. En este caso, es todavía más claro que no podemos decir que inferimos la creencia de que los niños están cometiendo un acto de crueldad sobre la base de la observación de su comportamiento, porque no creemos que estén cometiendo un acto de crueldad. Nos encontramos aquí con la tesis de la relativa independencia que, con respecto a las creencias, gozan las experiencias perceptivas.<sup>14</sup> Un sujeto que conoce el carácter ilusorio del caso de las líneas Müller-Lyer, que no cree, pues, que las líneas tienen diferente longitud, continúa experimentando las líneas como si tuvieran diferente longitud. De modo semejante, aun si se nos advierte que los niños que estamos observando están, en realidad, llevando a cabo una mera actuación, aun así, nos parece que están llevando a cabo un acto de crueldad. No importa aquí lo que creamos. Nuestra experiencia visual nos sigue mostrando un comportamiento que, fenoménicamente hablando, continúa pareciendo ser un caso típico de crueldad. Es justamente teniendo en cuenta este hecho que, al ejecutar su actuación, los actores pueden generar la impresión de que se está llevando a cabo un acto de crueldad. No se trata, pues, de que al ver a los niños adquirimos inmediatamente la creencia de que están cometiendo un acto de crueldad; antes bien, aunque creamos que se trata de una mera filmación, nos parece, fenoménicamente hablando, que se trata de un acto de crueldad. El que haya cierta fenomenología típica de los actos de crueldad que es independiente de las creencias o juicios que nos formemos al respecto (fenomenología que bien puede suscitar en nosotros impresiones o sentimientos asociados), constituye, me parece, un indicio sumamente sugerente de que, al menos en casos como el que estamos examinando, es posible experimentar las propiedades morales de los actos que las instancian.

¿Cómo es esto posible? Bueno, supongamos que el contenido de la experiencia perceptiva es de naturaleza conceptual; y supongamos también que conceptos como los de crueldad, valentía, justicia, etc., pueden formar parte del contenido de la experiencia moral. Tales conceptos están esencialmente conectados con el hacer y el elegir: su correcta aplicación depende de que se satisfagan ciertas condiciones objetivas, condiciones que involucran accio-

---

14 Véase Evans (1982).

nes realizadas intencionalmente por las personas en determinadas circunstancias.<sup>15</sup> Así, entendemos que la acción que llevan a cabo los niños del ejemplo de Harman satisface las condiciones de aplicación del término «cruel». El punto es que, si alguien posee el concepto de crueldad, no solo está en condiciones de juzgar que los niños están llevando a cabo un acto de crueldad, sino, además, de identificar perceptivamente que su acción es cruel. Puede ver, por ejemplo, que los niños, al hacer lo que hacen, están actuando intencionalmente, y que están causando dolor al gato innecesariamente. Puesto que la escena percibida cae bajo el concepto de crueldad, y —según hemos supuesto— este concepto puede formar parte del contenido de la experiencia, puede verse dicha escena —penetración cognitiva mediante— como una escena de crueldad. El patrón conductual ejemplificado por los niños en referencia al gato puede ser experimentado como un patrón que instancia la propiedad de la crueldad.<sup>16</sup>

Resulta crucial aquí resaltar el papel de la penetrabilidad cognitiva en la fenomenología de la experiencia.<sup>17</sup> La idea es, según reza la tesis, que ciertos estados mentales o capacidades (como conceptos, creencias, hipótesis, estados anímicos, emociones, deseos, rasgos de carácter) poseen el poder de afectar la fenomenología de la experiencia a punto tal que es posible que dos sujetos, por ejemplo, (o un sujeto en distintos momentos y circunstancias), podrían tener experiencias con fenomenologías diferentes aun cuando presten atención al mismo estímulo distal, bajo las mismas condiciones externas, como resultado de los distintos estados cognitivos o emocionales en los que se encuentran. Así pues, para volver a nuestro caso, quien posea el concepto de crueldad y, podemos agregar, sea sensible al dolor de otros seres vivos, eventualmente podrá experimentar, fenoménicamente hablando, la escena de los niños de Harman como una escena de crueldad.

---

15 Véase Foot (1994).

16 Se podría objetar aquí que solo podemos percibir, estrictamente, lo que concierne a la parte o dimensión descriptiva del concepto de crueldad. El aspecto valorativo que comporta este concepto quedaría, así, fuera del contenido perceptivo. Esta objeción parece presuponer que los conceptos densos, como el de crueldad, podrían ser descompuestos limpiamente en un elemento descriptivo y otro valorativo. Sin embargo, se ha argumentado convincentemente contra esta posibilidad. Véase McDowell (1998b, 1998c); Putnam (1988, cap. 9 y 2002, cap. 2). En cualquier caso, la idea es que, al figurar en el contenido de un estado experiencial, el concepto de crueldad nos capacita para ver no solo ciertos patrones conductuales, sino, además, para verlos con cierta carga valorativa negativa. Debo esta objeción a José Giromini.

17 Para la tesis de la penetrabilidad cognitiva, véase Siegel (2006, 2012), MacPherson (2012), Pylyshyn (1999), Stokes (2013), y Vance (2014).

Si todo esto es correcto, un juicio moral como, por ejemplo, «Esos niños están llevando a cabo un acto de crueldad con el gato», puede ser justificado apelando a la experiencia perceptiva correspondiente: juzgamos que se trata de un acto de crueldad porque *vemos* que se trata de un acto de tal naturaleza. Aunque en este caso la transición que va de la experiencia al juicio es inmediata y, según parece, no inferencial, no es preciso, desde luego, que todos los casos de justificación de un juicio por una experiencia perceptiva moral sean tan simples. La idea de que la experiencia moral pueda ser una razón para un juicio o creencia moral no excluye la posibilidad de que el vínculo de justificación sea, en casos más complejos, inferencial.

#### 4.

Hasta aquí he sugerido una explicación acerca de cómo conceptos densos como el de crueldad podrían figurar en el contenido de la experiencia moral, pero, ¿qué hay de términos delgados como «bueno», «malo», «correcto» o «incorrecto»? ¿También se puede decir que podemos ver que una acción posee alguna de esas propiedades? Considérese el caso de «correcto/a» o «incorrecto/a». Es obvio que estos términos tienen usos no morales. La gente suele decir, por ejemplo, que una movida de ajedrez es «correcta» cuando se ajusta a las reglas del ajedrez, que una palabra ha sido «incorrectamente» escrita cuando no se ajusta a las reglas ortográficas, o que «lo correcto» es devolver un saludo cuando uno ha sido saludado (algo que involucra ciertas reglas de cortesía).

Considérese ahora el caso en el que «correcto/a» tiene significado moral, o un uso moral. Según Ross (1930), dicho término significa «lo que es moralmente obligatorio» o «lo que es el deber hacer».<sup>18</sup> Lo que es moralmente obligatorio, o lo que se debe hacer, es lo que se conforma a la norma relevante. ¿Podemos percibir que la acción moral de una persona es correcta (o incorrecta) en este sentido? La tesis de la percepción moral sostiene que, en efecto, podemos percibir una acción como siendo correcta o incorrecta en la medida en que podamos percibirla como ajustándose o no ajustándose a la norma relevante.

Como en el caso de la crueldad recientemente analizado, el desafío consiste en responder al escéptico que sostiene que, literalmente hablando, nunca vemos acciones como ajustándose o no ajustándose a lo que debe hacerse. Lo

---

<sup>18</sup>El término en inglés es *Right*. Lo que es correcto hacer en ciertas circunstancias también se puede entender en un sentido prudencial: lo que es correcto en vistas de un fin. No me ocupo aquí de este sentido.

que siempre realmente hacemos, dice su objeción, es *juzgar*, sobre la base de lo que percibimos, que una acción moral se ajusta o no se ajusta a una norma.

No quisiera negar aquí que es posible, en efecto, juzgar que una acción es correcta o incorrecta sobre una base inferencial. Quizás esto suceda muchas veces en nuestra vida diaria. La idea sería, supongo, que la percepción (no moral) de cierta acción puede constituir una base inferencial para que, en conjunción con el conocimiento de la norma relevante, uno pueda juzgar que la acción en cuestión es o no es correcta según la norma en cuestión. Sin embargo, pienso que también es posible —y que muchas veces sucede de hecho— percibir acciones como moralmente correctas o incorrectas, o como ajustándose o no ajustándose a una norma moral.<sup>19</sup> También aquí se podría elaborar un experimento mental que, a semejanza del ya sugerido sobre los niños de Harman, pusiera en evidencia la independencia de la experiencia respecto del juicio o la creencia: podría suceder que, en una situación dada, no juzgáramos, o no estuviésemos dispuestos a juzgar, que cierta acción es incorrecta aun cuando de hecho nos pareciese, no obstante, que la acción es incorrecta.<sup>20</sup> Se sigue que, si una acción puede ser percibida como siendo incorrecta aun cuando no juzguemos ni creamos que en verdad lo es, entonces la percepción moral no colapsa con el juicio o la creencia moral. Por otro lado, como ha sido señalado por Audi (2013), hay veces en las que, al observar cierta acción, percibimos que algo «anda mal», aunque no podamos especificar en un juicio qué es exactamente lo que está mal con la acción percibida. En estos casos, dado el carácter difuso de la experiencia normativa, si bien percibimos, en un sentido vago, que la acción es incorrecta, no podemos juzgar que lo es, ni explicitar en qué sentido, o por qué, nos parece que es incorrecta. Por tanto, la percepción moral no es lo mismo que el juicio moral.

Con todo, es comprensible que persistan ciertas dudas acerca del carácter normativo o evaluativo de la percepción moral, pues, según todas las apariencias, no percibimos que una acción es correcta o incorrecta en el mismo sentido en que percibimos que una bufanda es roja o que cierto objeto es un libro. Además, la literatura sobre percepción moral no suele ser suficientemente explícita respecto de cómo es que podemos percibir acciones como siendo correctas o incorrectas desde un punto de vista moral. Pensemos, pues, en la percepción de una acción moral. Una acción en general es

---

19 Fuera del ámbito de la moral, también frecuentemente decimos que vemos que una palabra está incorrectamente escrita, que un jugador de fútbol ha cometido una infracción, o que una pieza de ajedrez ha sido movida incorrectamente.

20 Por ejemplo, pese a haber visto hacer algo absolutamente incorrecto a alguien muy cercano a mí, podría ser renuente a darle crédito a mi experiencia y, por ende, no creer que ha hecho lo que he visto que ha hecho.

un cierto tipo de evento. Los eventos son entidades que tienen cierta duración. Percibirlos, pues, no es como percibir un objeto material. Para percibir un evento tenemos que esperar a que ocurra y, una vez iniciado, debemos esperar a que se despliegue en el tiempo. No obstante, aunque esto sea así, frecuentemente somos capaces de identificar un evento cuando solo hemos podido percibir una parte de él. Es por esto que, junto con la percepción de los inicios de un evento, por ejemplo, podemos tener ciertas expectativas acerca de cómo se irá desarrollando en el futuro inmediato. Ahora bien, aquí, se podría pensar en una extensión de la tesis de la percepción a-modal. Como se recordará, esta tesis decía que percibimos a-modalmente la parte oculta de un objeto cuando percibimos su cara frontal. Esta idea quizá pueda aplicarse también a la percepción de eventos, especialmente cuando son de corta duración. Podría decirse: al percibir los inicios de cierto evento, nos representamos a-modalmente las otras partes del evento que más o menos inmediatamente tendrán lugar. La percepción del despliegue completo de tal evento corroborará o falsará las expectativas generadas por la percepción a-modal. Desde luego, podría suceder que no fuera correcto extender la idea de la percepción a-modal al caso de los eventos, como estoy sugiriendo. Quizás, todo lo que sucede es que, al percibir un momento del despliegue de un evento que hemos identificado, simplemente nos formamos inferencialmente ciertas expectativas acerca de lo que sucederá a continuación. En cualquier caso, lo importante es esto: que, al percibir parte de un evento, podemos identificarlo como el evento que es, y que la percepción de parte del evento genera expectativas acerca de lo que acontecerá a continuación.

Pues bien, se podría pensar ahora que algo semejante ocurre en el caso de la percepción moral. Al observar los inicios de una acción moral, y dada nuestra comprensión de lo que debería hacer *S*, el agente moral que estamos observando, en su situación, nos representamos a-modalmente, o simplemente llegamos a pensar o esperar, no solo lo que probablemente ocurrirá, sino lo que *debería* ocurrir si es que *S* hiciera lo que es debido en esa situación. Si *S* hace lo que debe hacer, entonces la percepción de cómo se despliega su acción satisfará nuestras expectativas normativas; en caso contrario, las defraudará. En el primer caso, se podría decir entonces que percibimos la acción de *S* como ajustándose a la norma relevante; en el segundo, que percibimos la acción de *S* como no ajustándose a ella. A diferencia de lo que ocurre con la percepción de un evento natural, el que *S* no satisfaga nuestras expectativas generadas por la percepción a-modal de las partes o aspectos de su acción todavía no ejecutados no significa que estemos equivocados. Todo lo contrario, significa, si estamos en lo correcto acerca de lo que debía hacer *S*, que *S* ha hecho algo incorrecto.



En este punto, resulta crucial advertir que estamos suponiendo la tesis de la penetrabilidad cognitiva: es porque las normas morales penetran en la experiencia perceptiva que, al observar la conducta de *S*, puede parecernos, fenoménicamente, que ella es correcta o incorrecta. El tener conocimiento de una norma como la que dice, por ejemplo, «No se debe hacer *F*», nos permite identificar perceptivamente una acción, que cuenta como un caso de hacer *F*, como una acción que se ajusta o no se ajusta a dicha norma. No es preciso postular aquí alguna suerte de percepción de la corrección o incorrección mismas, como si se tratara de propiedades que «acompañan» a las acciones morales.<sup>21</sup> Percibir una acción como correcta o incorrecta es, simplemente, percibirla como ajustándose o no ajustándose a la norma relevante.

Considérese ahora el término «bueno». ¿Podemos percibir las acciones morales como siendo buenas (o malas)? Las dificultades surgen, aquí, no debido al hecho de que «bueno» tenga usos no morales (v.g., «un buen cuchillo»), ni a que su contenido descriptivo (si es que lo tiene) sea mínimo, sino, fundamentalmente, a que los filósofos no están de acuerdo en absoluto acerca del significado de «bueno». Como se sabe, hace tiempo Moore sostuvo que «bueno» era un término indefinible, que aludía a una propiedad simple no natural (Moore, 1903). Por su parte, Ross (1930) sostuvo que el término «bueno», en su uso moral, es ambiguo. Puede significar (i) «que conduce al fin deseado»; (ii) «intrínsecamente bueno» (esto es, bueno independientemente del resultado que produzca); (iii) «contributivamente bueno» (que agrega cierto valor a un todo que es intrínsecamente bueno); o (iv) «lo que es últimamente bueno» (bueno por sí mismo y no por los efectos que produce o por los elementos que contiene). Sin embargo, varios filósofos han sido escépticos acerca de la idea de que hay algo así como «ser bueno *simpliciter*» (Geach, 1956; Thompson, 1994; Foot, 1985). De acuerdo a tales autores, las adscripciones de bondad son, o bien todas conceptualmente reducibles a otras formas de bondad, o bien carecen de sentido.<sup>22</sup> Con todo, cualquiera sea el resultado de esta disputa, parece que el defensor de la tesis de la percepción moral puede pergeñar una estrategia que lo mantenga al margen de esta disputa semántica. Puede sostener que, si «bueno» en última instancia refiere, o es reducible a, distintas formas de bondad (v.g., valentía, justicia,

---

21 Para referirse a la relación entre las propiedades morales y las propiedades físicas en las que las primeras supuestamente se instancian, los filósofos que defienden la tesis de la percepción moral suelen hablar de «superveniencia» (las propiedades morales supervienen a las propiedades físicas), de «fundamentación» (las propiedades físicas «sostienen», «fundamentan» o «soportan» a las propiedades morales), o de una relación «consecuencial» (las propiedades morales son «consecuentes» de las propiedades físicas). No puedo discutir este importante problema aquí. Al respecto, véase Audi (2013), Blackburn (1988), McDowell (1988) y Boyd (1988).

22 Para una defensa reciente de lo bueno *simpliciter*, véase Rowland (2015).

honestidad, sinceridad, etc.), entonces, su tesis acerca de la percepción de ciertos hechos morales no resulta afectada por el debate semántico, pues ya se ha «probado» que en tales casos resulta plausible afirmar que percibimos, en efecto, acciones valientes, justas o sinceras (así como, según se argumentó anteriormente, percibimos actos de crueldad). Si, en cambio, resultara que «bueno» alude a una propiedad irreductible, a lo bueno *simpliciter*, entonces podría argumentar que, dependiendo del sentido de «bueno» que esté en juego en cada caso, podrían ser percibidas acciones que son buenas para alcanzar algún fin deseado, o que son buenas en sí mismas, independientemente de los resultados que permitan alcanzar. Una vez que se han asumido (i) la tesis conceptualista acerca del contenido de la experiencia moral y (ii) la tesis de la penetrabilidad cognitiva, y (iii) la tesis de la percepción a-modal, no se vislumbran —puede señalar— dificultades específicas para con el término «bueno».

## 5.

En este artículo, he tratado de presentar la mejor versión posible de la tesis de la percepción moral, haciendo explícitos los compromisos teóricos involucrados en ella. El argumento presentado en § 4, que apela al carácter fenoménico de la experiencia moral, hace plausible (junto con el argumento del contraste fenoménico, presente en la literatura), la tesis de la percepción moral. Las distinciones posteriores han sido introducidas para explicitar lo que está involucrado en dicha tesis. Así, he sostenido que la experiencia moral ha de ser concebida como teniendo carácter representacional, carácter este que incluye, fundamentalmente, la representación de propiedades de alto nivel como las propiedades morales. He dicho también que el contenido de la experiencia moral ha de ser pensado como siendo de naturaleza conceptual, lo cual involucra tanto conceptos densos como delgados. Asimismo, he indicado que la percepción a-modal podría desempeñar un papel central en la percepción de las acciones morales. Si este cuadro de lo que es la percepción de hechos morales es mínimamente correcto, puede comenzar a vislumbrarse cómo la percepción moral podría tener, a semejanza de la experiencia empírica, un papel epistémico en el conocimiento moral. Por ejemplo, el juicio de que *S* ha cometido una injusticia con *R* podría ser justificado por la percepción de la acción de *S* (que es percibida como injusta). Asimismo, la creencia de que *S* ha obrado incorrectamente (desde el punto de vista moral) con respecto a *R* podría ser verificada por la percepción de la acción de *S*. Y el juicio de que *S* ha obrado bien en tales y cuales circunstancias, o de que ha

hecho algo que es intrínsecamente bueno, también podría estar justificado en la percepción de lo que *S* ha hecho. En todos estos casos, percibir lo que *S* ha hecho no es meramente percibir ciertas acciones en determinadas circunstancias, sino, además, percibir tales acciones como instanciando ciertas cualidades morales. De esta forma, la experiencia puede justificar, inferencial o no inferencialmente, nuestros juicios y creencias morales y, así, proveer a estos de un anclaje en la realidad moral.

Con todo, la tesis de la percepción moral, tal como la he especificado, posee sus limitaciones epistémicas. En efecto, puesto que, según se ha sostenido, el contenido de la percepción moral es de carácter conceptual, la percepción moral misma no puede brindarnos una ayuda directa en la justificación de tales conceptos y normas. En tal sentido, solo puede revelarnos que las acciones que percibimos caen bajo tales conceptos y normas y, así, puede proveernos conocimiento perceptivo acerca del carácter moral de tales acciones; sin embargo, puesto que en su mismo contenido la percepción simplemente presupone tales conceptos y normas, no puede ayudarnos a evaluar si ellos son los conceptos y normas que *deberíamos* tener. En este sentido, aunque podamos ver, por ejemplo, que cierta acción se ajusta o no se ajusta a una norma dada, ello no puede ayudarnos a determinar si la norma misma es correcta o no.<sup>23</sup>

Pese a esta limitación, quizás la percepción moral todavía podría ser de ayuda, aunque de un modo indirecto, en la empresa de evaluar reflexivamente si cierta norma es realmente correcta o no, a saber, revelándonos si su aplicación en distintas ocasiones y circunstancias produce resultados justos, crueles, etc. En un caso así, estaríamos presuponiendo que los conceptos densos que figuran en el contenido de la experiencia pueden ayudarnos a determinar si una norma particular es correcta. Así, por ejemplo, una norma como «No se debe mentir» podría tener cierta justificación en la percepción de los resultados que produce su seguimiento generalizado. Naturalmente, esto último dependerá, a su vez, de nuestros conceptos de lo que es dañoso, perjudicial, beneficioso, etc. En cualquier caso, enfrentamos aquí el difícil problema de la justificación de las normas y valores morales, un problema que, seguramente, no puede ser desarrollado aquí.<sup>24</sup>

---

23 Estoy aplicando aquí, para el caso de la percepción moral, una objeción originalmente presentada por Siegel (2012) con relación a la capacidad justificatoria de la percepción cuando se asume la tesis de penetrabilidad cognitiva.

24 Agradezco profundamente a Laura Danón y a José Giromini por sus agudas observaciones a una versión previa de este artículo.

## Referencias bibliográficas

- Aristóteles (2000). *Ética Nicomaquea* (trad. de Julio Pallí Bonet). Gredos.
- Audi, R. (2004). *The Good in the Right*. Princeton University Press.
- Audi, R. (2010). Moral Perception and Moral Knowledge. *Proceedings of the Aristotelian Society, Supplementary Volume LXXXIV*, (1), 79-97.
- Audi, R. (2013). *Moral Perception*. Princeton University Press.
- Audi, R. (2015). Moral Perception Defended. *Argumenta*, 1(1), 5-28.
- Austin, J. (1962). *Sense and Sensibilia*. Oxford University Press.
- Bayne, T. (2010). Perception and the Reach of Phenomenal Content. *The Philosophical Quarterly*, 59(236), 385-404.
- Blackburn, S. (1988). Supervenience Revisited. En Sayre-McCord, G. (Ed.), *Essays on Moral Realism* (pp. 59-75). Cornell University Press.
- Bloom, L. (1994). *Moral Perception and Particularity*. Cambridge University Press.
- Boyd, R. (1988). How to Be a Moral Realist. En Sayre-McCord, G. (Ed.), *Essays on Moral Realism* (pp. 181-228). Cornell University Press.
- Brandom, R. (1994). *Making it Explicit*. Harvard University Press.
- Brewer, B. (1999). *Perception and Reason*. Oxford University Press.
- Brewer, B. (2006). Perception and Content. *European Journal of Philosophy*, 14, 165-181.
- Brewer, B. (2008). How to Account for Illusion. En Haddock, A. y Macpherson, F. (Eds.), *Disjunctivism: Perception, Action, Knowledge* (pp. 168-180). Oxford University Press.
- Brewer, B. (2011). *Perception and Its Objects*. Oxford University Press.
- Briscoe, R. (2011). Mental Imagery and the Varieties of Amodal Perception. *Pacific Philosophical Quarterly*, 92(2), 153-173.
- Campbell, J. (2002). *Reference and Consciousness*. Oxford University Press.
- Cowan, R. (2015). Perceptual Intuitionism. *Philosophy and Phenomenological Research*, 90(1), 164-193.
- Crane, T. (1992). The Nonconceptual Content of Experience. En Crane, T. (Ed.), *The Contents of Experience* (pp. 136-157). Cambridge University Press.
- Cullison, A. (2009). Moral Perception. *European Journal of Philosophy*, 12(2), 159-175.
- Cussins, A. (2003). Content, Conceptual Content, and Nonconceptual Content. En Gunther, Y. (Ed.), *Essays on Nonconceptual Content* (pp. 133-163). The MIT Press.
- Chappell, T. (2008). Moral Perception. *Philosophy*, 83(4), 421-437.
- Dancy, J. (2010). Moral Perception. *Proceedings of the Aristotelian Society* (Supplementary Volume), LXXXIV(1), 99-117.
- Davidson, D. (2001). Rational Animals. En *Subjective, Intersubjective, Objective* (pp. 95-105). Clarendon Press.
- Dretske, F. (1999). *Knowledge and the Flow of Information*. The MIT Press.
- Evans, G. (1982). *Varieties of Reference*. Oxford University Press.
- Fish, W. (2009). *Perception, Hallucination, and Illusion*. Oxford University Press.
- Foot, Ph. (1985). Utilitarianism and the Virtues. *Mind*, 94(374), 196-209.
- Foot, Ph. (1994). Argumentos morales. En *Las virtudes y los vicios* (pp. 117-131). UNAM (México).
- Geach, P. (1956). Good and Evil. *Analysis*, 17(2), 33-42.

- Greco, J. (2000). *Putting Skeptics in Their Place*. Cambridge University Press.
- Harman, G. (1977). *The Nature of Morality*. Oxford University Press.
- Heck, R. (2000). Nonconceptual Content and the «Space of Reasons». *The Philosophical Review*, 109(4), 483-523.
- Hinton, J. (1973). *Experiences*. Clarendon Press.
- Huemer, M. (2008). *Ethical Intuitionism*. Palgrave Macmillan.
- Hume, D. (1738/1984). *Tratado de la naturaleza humana* (trad. de Félix Duque). Hyspamérica.
- Johnston, M. (2006). Better than Mere Knowledge? The Function of Sensory Awareness. En Gendler, T.Z. y Hawthorne, J. (Eds.), *Perceptual Experience* (pp. 260-290). Clarendon Press.
- Kalpokas, D. (2018). Non-conceptualism and the Myth of the Given. *Dialectica*, 72(3), 331-363.
- Kirchin, S. (Comp.) (2013). *Thick Concepts*. Oxford University Press.
- MacPherson, F. (2012). Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism. *Philosophy and Phenomenological Research*, 84(1), 24-62.
- Mackie, J. (1977). *Ethics: Inventing Right and Wrong*. Penguin Books.
- Martin, M. (2002). The Transparency of Experience. *Mind & Language*, 17(4), 376-425.
- Martin, M. (2009a). The Reality of Appearances. En Byrne, A. y Logue, H. (Eds.), *Disjunctivism* (pp. 91-116). The MIT Press.
- Martin, M. (2009b). The Limits of Self-Awareness. En Byrne, A. y Logue, H. (Eds.), *Disjunctivism* (pp. 271-317). The MIT Press.
- McBrayer, J. (2010). A limited Defense of Moral Perception. *Philosophical Studies*, 149(3), 305-320.
- McGrath, S. (2004). Moral Knowledge by Perception. *Philosophical Perspectives*, 18(1), 209-228.
- McDowell, J. (1988). Values and Secondary Qualities. En Sayre-McCord, G. (Ed.), *Essays on Moral Realism* (pp. 166-180). Cornell University Press.
- McDowell, J. (1996). *Mind and World*. Harvard University Press.
- McDowell, J. (1998a). Virtue and Reason. En *Mind, Value, & Reality* (pp. 50-73). Harvard University Press.
- McDowell, J. (1998b). Non-Cognitivism and Rule-Following. En *Mind, Value, & Reality* (pp. 198-218). Harvard University Press.
- McDowell, J. (1998c). Aesthetic Value, Objectivity, and the Fabric of the World. En *Mind, Value, & Reality* (pp. 112-130). Harvard University Press.
- Moore, G. (1903/1997). *Principia Ethica* (trad. Adolfo García Díaz y Ana Isabel Stellino). UNAM.
- Nanay, B. (2010). Perception and Imagination: Amodal Perception as Mental Imagery. *Philosophical Studies*, 150(2), 239-254.
- Nanay, B. (2013). *Between Perception and Action*. Oxford University Press.
- Noë, A. (2004). *Action in Perception*. The MIT Press.
- Peacocke, C. (1992). *A Study of Concepts*. The MIT Press.
- Putnam, H. (1988). *Razón, verdad e historia*. Tecnos.
- Putnam, H. (2002). *The Collapse of the Fact/Value Dichotomy*. Harvard University Press.

- Plyshyn, Z. (1999). Is Vision Continuous with Cognition? The Case for Cognitive Impenetrability of Visual Perception. *Behavioral and Brain Sciences*, 22(3), 341-365.
- Raleigh, Th. (2013). Phenomenology Without Representation. *European Journal of Philosophy*, 21(3), 1209-1237.
- Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton University Press.
- Rorty, R. (1998). The Very Idea of Human Answerability to the World: John McDowell's Version of Empiricism. En *Truth and Progress* (pp. 138-152). Cambridge University Press.
- Ross, W. (1930). *The Right and the Good*. Oxford University Press.
- Rowland, R. (2016). In Defense of Good Simpliciter. *Philosophical Studies*, 173(5), págs. 1371-1391.
- Ryle, G. (1949). *The Concept of Mind*. Hutchinson's University Library.
- Searle, J. (1983). *Intentionality*. Cambridge University Press.
- Siegel, S. (2006). Which Properties Are Represented in Perception? En Gendler, T.Z. y Hawthorne, J. (Eds.), *Perceptual Experience* (pp. 481-503). Clarendon Press.
- Siegel, S. (2010). *The Contents of Visual Experience*. Oxford University Press.
- Siegel, S. (2012). Cognitive Penetrability and Perceptual Justification. *Nous*, 46(2), 201-222.
- Stokes, D. (2013). Cognitive Penetrability of Perception. *Philosophy Compass*, 8(7), 646-663.
- Stratton-Lake, Ph. (2003). *Ethical Intuitionism*. Clarendon Press.
- Thomson, J. (1994). Goodness and Utilitarianism. *Proceedings and Addresses of the American Philosophical Association*, 67(4), 145-159.
- Travis, Ch. (2004). The Silence of the Senses. *Mind* 113(449), 57-94.
- Tye, M. (1995). *Ten Problems of Consciousness*. MIT Press.
- Vance, J. (2014). Emotion and the New Epistemic Challenge from Cognitive Penetrability. *Philosophical Studies*, 169(2), 257-283.
- Werner, P. (2016). Moral Perception and the Contents of the Experience. *Journal of Moral Philosophy*, 13(3), 294-317.
- Williams, B. (1985). *La ética y los límites de la filosofía*. Monte Ávila Editores.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.
- Wisniewski, J. (2015). The Case for Moral Perception. *Phenomenology and the Cognitive Sciences*, 14(1), 129-148.
- Woods, M. (1986). Intuition and Perception in Aristotle's Ethics. En Woods, M. (Ed.), *Oxford Studies in Ancient Philosophy* (pp. 145-166). Vol. 4. Clarendon Press.

## 10. Argumentación moral

*Daniel Busdygan*

La argumentación constituye una actividad racional a través de la cual se producen argumentos con el propósito de probar o refutar una tesis a partir del apoyo fundado en razones. La finalidad de la argumentación está orientada a la persuasión racional de otro/s agente/s quien/es puede/n llegar a la comprensión o el asentimiento sobre la afirmación que se presenta como tesis del tema que está en cuestión (Vega Reñón, 2011:66). La racionalidad, así entendida, consiste en el atributo que tienen las personas para adquirir, formar o cambiar sus creencias a partir de la evidencia disponible, la cual pueden evaluar de formas intersubjetivamente válidas con precisión y cuidado, evitando sesgos, fallas producto de la motivación o modos inadecuados de razonar. Cuando razonamos sobre algún tema, producimos argumentos donde brindamos —exponemos— ciertas afirmaciones en apoyo o sustento de otra: «una afirmación es, primariamente, un decir algo acerca de algo; y es un acto sujeto a condiciones de aceptabilidad intersubjetiva» (Moretti, 2020:525).

Un argumento, en tanto unidad lingüística particular que produce la argumentación, es un conjunto de proposiciones donde se realiza y se detecta un proceso inferencial. El argumento es así un segmento lingüístico que puede ser desagregado de la argumentación para ser analizado y evaluado separadamente de modo externo e interno (Plantin y Muñoz, 2011:83). Para esto último es preciso llevar adelante un proceso de análisis, interpretación y reconstrucción —sin merma cognoscitiva— del argumento, para obtener un esquema sintáctico que *exhiba* el modo de razonar, su forma lógica y la pertinencia que se advierte en la conexión inferencial detectada; ulteriormente, la evaluación se lleva a cabo conforme a parámetros propios de la lógica formal o de una teoría de la argumentación. El proceso de reconstrucción argumentativa se orienta a la aclaración de su estructura para mejorar no solo la intelección de lo que se afirma y sus implicancias sino además el modo en el que se presentan las razones que se proveen y el tipo de las mismas. Es por ello que en la actividad reconstructiva, no se realiza solo una tarea descriptiva, pues allí se pone de relevancia que «hay prácticas inferenciales. La práctica inferencial es evaluativa, incluye juicios de valor acerca de argumentos, por tanto es normativa» (Moretti, 2020:531).

En la estructura de todo argumento se pueden reconocer tres elementos esquematizables: premisa/s, conclusión —ambas partes con diferente función y estatus— y la conexión inferencial. Por una parte, el conjunto de *premi-*

*sas* está comprendido por las afirmaciones —o hipótesis— que operan como base, apoyo, sustento, o razón/es para que, proceso inferencial mediante, se siga la derivación de una afirmación que constituirá la *conclusión*. En su etimología latina, la palabra «premisa» proviene de *praemisa* que significa «puesta o colocada delante» (Ausín y Peña, 2011:479) y designa figurativamente muy bien esa función que cumple aquel grupo de proposiciones dentro del razonamiento. Por otra parte, la «conclusión» es el «cierre» de la inferencia —del latín, *conclusio, claudere*— y da cuenta de aquello que el argumento pretende establecer como tesis dentro del tema en cuestión.

Las premisas buscan conducir al/los destinatario/s del argumento hacia la conclusión y, efectivamente, eso puede darse a través de distintas formas de argumentar, deductivas o no deductivas. En el primer caso, las premisas permiten que la conclusión se siga como una consecuencia lógica, con necesidad absoluta, con independencia del agregado de premisas adicionales (monotonía) y sin la admisión de grados de probabilidad (Legris, 2001; Copi, 2009; Gamut, 2009). Por otro lado, en el caso de las inferencias no deductivas, si bien las premisas pueden proveer buenas razones para que siga la conclusión, las mismas no son suficientes lo cual implica que la conclusión esté sujeta a ciertos grados de probabilidad y posea una naturaleza revocable —retractable—, pues la misma podría tener que revisarse en caso que se incorporen nuevas premisas. Las formas no deductivas pueden presentar argumentos convincentes al momento de revisar la evidencia que guía nuestras creencias, y esto se puede dar de varias formas en la conexión inferencial: inducción, analogía, abducción, inferencia rebatible, entre otras (Copi, 2009; Strawson, 2011). En este punto, y siguiendo a Moretti, es importante apuntar:

la distinción entre contexto de emisión y contexto de evaluación de razonamientos, que permite variar las pretensiones de fuerza inferencial (deducciones, inducciones, abducciones) que se adscriben a un argumento en el momento de su evaluación. Los argumentos *per se* no son deductivos o no deductivos; lo son nuestras pretensiones al emitirlos y nuestros criterios al evaluarlos. (2020:539)

En lo que sigue será por demás importante revisar algunas cuestiones teóricas vinculadas con los *argumentos morales* y los procesos de reconstrucción para la evaluación de los mismos. En tal sentido, veremos aspectos relevantes al momento de evaluar el peso que pueden tener tanto las razones como el modo en el que se realizan las inferencias a partir de esas razones. Comencemos dándonos una definición amplia de estos. Los argumentos morales están enlazados con el mundo moral en tanto nos hablan de lo que es bueno/malo, correcto/correcto, moral/inmoral, o lo que debería/no debería hacerse.



Se argumenta moralmente en un sinnúmero de asuntos significativos para nuestra vida en comunidad. Ejemplo de ello es cuando se sostiene que debería/ no debería permitirse el uso recreativo de drogas, la tauromaquia, el aborto, la eutanasia, el aumento progresivo de la base impositiva, la tenencia libre de armas de fuego, la pornografía, el impuesto a la herencia, la pena de muerte, la investigación en animales, etc. Mediante la argumentación moral se exponen, tanto para otros como para nosotros mismos, cuáles son las razones que darían sustento —y qué tipo de sustento darían, sólido, fuerte, débil— para que se siguiesen ciertas acciones, actitudes o juicios morales. Es así que en estos argumentos, la aceptación que se pretende establecer de la conclusión no aspira a convertirse en una buena pieza de teórica sino que pretende que dicha aceptación se corresponda con algún tipo de compromiso —en algún grado considerable— coherente con el contenido de esa creencia o juicio y que, evidentemente, se traduzca en el comportamiento.

En términos muy generales, cualquier argumento moral puede ser caracterizado con la misma estructura con la que se representa frecuentemente a los otros tipos de argumentos, a saber, formados por premisas y conclusión. No obstante, si bien son comparables con otros argumentos que responden a otros campos de la argumentación, poseen aspectos específicos. Por un lado, a diferencia de los razonamientos teóricos donde el conjunto de proposiciones está compuesto solo por proposiciones descriptivas, en estos razonamientos se incorporan proposiciones morales que pueden ser evaluativas o prescriptivas. Por otro lado, si en otros razonamientos el conjunto de premisas puede estar compuesto extensionalmente por una única premisa, en el caso de los argumentos morales eso no es posible, pues para su conformación como razonamiento bien formado se requieren al menos dos premisas, las cuales deben ser distintas: una premisa moral (en adelante,  $\text{PM}$ ) y otra premisa no-moral (en adelante,  $\text{P}\leftarrow\text{M}$ ). De aquí que: si  $\phi$  es un argumento moral, entonces su conclusión es una proposición moral que fue obtenida a partir de un conjunto finito de premisas  $\Gamma$ . Asimismo, el conjunto  $\Gamma$  está compuesto por dos grupos de proposiciones de una naturaleza distinta, un grupo que comprende aspectos referidos a la moral y otro que refiere a temas no morales. Por último, hay al menos una premisa explícita (en adelante,  $\text{PE}$ ) o implícita (en adelante,  $\text{PI}$ ) en cada grupo.

Si  $\phi$  es un argumento moral, entonces su estructura es  $p_1, \dots, p_2, \dots \vdash C$   
 $p_1 \in \Gamma$  y es una o más premisa con contenido moral ( $\text{PM}$ ). Explícita o implícitas ( $\text{PE} \vee \text{PI}$ )

$p_2 \in \Gamma$  y es una o más premisa con contenido no moral ( $P \rightarrow M$ ). Explícita o implícitas ( $PE \vee PI$ )

C se sigue de  $\Gamma$  y es una proposición con contenido moral

Sumariamente, algunas consideraciones sobre los dos tipos de proposiciones que entran en relación con un argumento moral. Por un lado, las proposiciones descriptivas constituyen una unidad de información que da cuenta de un estado de cosas empíricamente comprobable del mundo. Es por eso por lo que de ellas tiene sentido decir desde el metalenguaje valorativo que son verdaderas o falsas, es decir, poseen un valor de verdad inscripto en una relación correspondentista entre lenguaje y mundo. En este sentido, una proposición verdadera es aquella que se ajusta a lo que es en el mundo. Por otro lado, dentro de las proposiciones morales podemos diferenciar dos tipos de proposiciones: las prescriptivas y las evaluativas o valorativas.

Las proposiciones prescriptivas ordenan que se realice o deje de realizar una acción. La relación entre lenguaje y mundo se encuentra en un sentido contrario al anterior, el lenguaje aquí busca procurar que se dé un determinado estado de cosas en el mundo. El sentido de ajuste es el inverso (Guariglia y Vidiella, 2011:30). Las proposiciones prescriptivas poseen un núcleo descriptivo el cual está referido a la acción que se ordena. Esas proposiciones tienen dos posibilidades al momento de ser evaluadas desde el metalenguaje, son correctas o incorrectas. A diferencia de estas, las proposiciones evaluativas incorporan un término evaluativo que puede concebir gradientes y se halla en función de las preferencias del hablante —el ejemplo recurrente es «x es bueno/a»—. En la expresión utilizada en estas proposiciones se introduce adjetivación que denota una apreciación positiva o negativa —bueno o malo—, comparativa —mejor o peor—, o superlativa —óptimo o pésimo.

Las características de las proposiciones morales incorporan particularidades para la evaluación de argumentos. Y muchas de esas particularidades se muestran como dificultades vinculadas con la verdad, la validez, la aceptabilidad y la racionalidad de estos enunciados. En tal sentido, cuando se trabaja con estos argumentos, debe tenerse en consideración las cuestiones vinculadas con el significado de estas proposiciones y las diferencias que existen entre sus valores metalingüísticos y la relación ontológica que se habilita entre las proposiciones y el mundo.

Para muchos, las proposiciones morales constituyen todo un problema si se considera el estatus ontológico de las entidades a las que hacen referencia. En correspondencia con esto se sostiene que la alusión a determinados términos termina por volver irresolubles los problemas morales para personas razonables y bien dispuestas a considerar todos los puntos de una disputa —es decir,

excluidos los necios y fanáticos para quienes ninguna evidencia es suficiente (Feldman, 1999)—. Que pueden ser irresolubles, se sigue de que la evidencia nunca termina por ser concluyente en algunos temas o porque no se puede disponer de ella con algún grado aceptable de objetividad. Este hecho afecta al estatus de estas proposiciones. Para otros, en cambio, dado que toda moral es relativa a su contexto, ello les imprime a las proposiciones morales un peso relativo o, aún más, hasta podrían ser tomadas como consideraciones subjetivas —especialmente las valorativas—. Ahora bien, no avanzaremos en estas discusiones dado que las mismas se apartan de los propósitos del presente capítulo. No obstante, muy sucintamente digamos que si bien es cierto que se puede ser relativista o escéptico en relación con muchas cuestiones morales, de eso no sigue que estemos parados sobre un terreno de indefiniciones absolutas. Considero que no sería inapropiado decir que algunas proposiciones morales están lejos de ser temas irresueltos, abiertos, sujetos al contexto o a la subjetividad; por ejemplo, sostener que «torturar a un niño/a por placer está mal» es una proposición moral discutible no parece una opción. Así, existen otros ejemplos análogos que podríamos encontrar relacionados con cuestiones tales como la esclavitud, la dignidad de las personas, la discriminación, entre otros puntos (Feldman, 1999:377; Busdygan, 2017).

Los argumentos morales pueden ser reconstruidos sobre la base de procedimientos que permitan evaluar aspectos vinculados con su estructura. En tal sentido, al momento de ser reconstruida la proposición moral es importante evitar cualquier tipo de *vaguedad*, esto es, que su contenido sea impreciso o indefinido y no permita establecer parámetros de aplicación. Así también debe tenerse en cuenta la supresión de cualquier tipo de *ambigüedad* donde pueda leerse más de un significado en la proposición. Y por último, debe evitarse que las proposiciones morales se presenten en su reconstrucción como *incompletas* o que en ellas se omitan expresiones lógicas cuantificadoras —todos/algunos— u operadores modales de temporalidad —siempre/nunca— en caso que sean necesarios.

Luego de establecer un primer acercamiento general, establezcamos otro punto que describe a los argumentos morales y que deberíamos poder observar en sus esquemas. En los razonamientos morales pueden diferenciarse proposiciones que aluden a *tres niveles distintos* dentro de sus esquemas formales. Cada nivel está complementado con los otros para la derivación de la conclusión. Asimismo, al momento del análisis y la reconstrucción, dichos niveles pueden desagregarse y están abiertos o son factibles de cuestionamientos críticos que pueden realizarse en caso de que se requiera mayor explicación, explicitación, clarificación o reafirmación en lo que allí se afirme o se presuma. Los tres niveles en cuestión son el *nivel epistémico*, el *nivel moral* y otro,

superior, el *nivel de los principios* (Guariglia y Vidiella, 2011). Mientras que el nivel epistémico y el moral buscan establecer un conjunto de certidumbres compartidas, el nivel de los principios suele estar supuesto o contenido en la norma a la que refiere el nivel moral (Guariglia, 1996:53).

El *nivel epistémico* nos presenta información referida a datos y/o acciones vinculadas con hechos empíricos que se reportan en el razonamiento particular. Sea de modo explícito o implícito (entimemas) es necesario que en un argumento moral se dé al menos una premisa cuyo contenido no sea moral, pues, de lo contrario no sería posible obtener una conclusión moral sin que esté dada esa premisa (Nunes y Salvadori, 2020). Para que se acredite el carácter público de una proposición descriptiva y se logre un cierto «suelo común» entre hablantes racionales y razonables, las aserciones que se trazan allí deberían ser directa o indirectamente *factibles* de comprobación empírica —es decir, «La vacuna *a* reportó un caso de embolia cada cien mil habitantes» o «La Tierra es plana»—. Ahora bien, deben diferenciarse y excluirse a las *pseudo* proposiciones de las que sí lo son pero que, por cuales sean las razones, se dificulta el acceso empírico a evidencia concluyente y desconocemos su valor de verdad —la paleontología y la astronomía son buenos ejemplos de disciplinas con estas dificultades—; es decir, que no sepamos si una proposición es verdadera o falsa simplemente no implica que no posee un valor verdad. Bien, como veremos seguidamente, la premisa no moral es *necesaria* dentro de la estructura del esquema de un argumento moral.

En la premisa con contenido moral y en la conclusión aparece el *nivel moral*. Allí se establece una relación entre la evaluación o la prescripción moral de una acción o de una clase de acciones y el contenido de la premisa no moral. La premisa moral incorpora el conocimiento moral en la inscripción que se hace de la acción a una norma y lo anticipa a la conclusión, pues ella está dada dentro de ese rango operativo. La acción en cuestión posee entonces las características típicas para poder ser incluida y juzgada a partir de un criterio normativo o valorativo que se proporciona para evaluar las acciones (Guariglia, 1996:151). En tal sentido, es necesario explicitar o suponer una norma válida para establecer la corrección o incorrección de una acción.

Para convenir con otro/s sobre un juicio moral y trazar aquí ese «suelo común», se precisa necesariamente, por un lado, suponer alguna norma general válida que se ha cumplido o se ha infringido, y por otro, hacerlo desde el punto de vista de la tercera persona. En relación con lo primero, podemos aceptar la conclusión «Pedro actuó incorrectamente al negar que conocía a Jesús» solo si dicha acción tipifica como una acción contraria a lo que estipula una norma, en este caso moral, la cual debe explicitarse y donde se establece qué es lo correcto. Lo que justificaría dicha inferencia es que la misma se haya realizado

asumiendo que «Nunca se debe desconocer a ningún amigo». Las acciones o se adecuan a normas generales que se enuncian o no lo hacen y las lesionan. De este modo, los hechos pueden darse por válidos en función de los aspectos veritativos propios del lenguaje descripto al mismo tiempo que las acciones son correctas o no bajo la pretensión de validez del lenguaje moral.

Ahora bien, por otro lado, la conclusión en el argumento debe ser una proposición dada desde el punto de vista de un observador y esto con el propósito de satisfacer la requerida objetividad del juicio moral. La conclusión de un argumento moral no tiene por propósito dar cuenta de un estado de cosas sino que constituye una fuerza ilocucionaria que busca provocar un determinado estado de cosas a través del cumplimiento de un imperativo o de una norma. No es lo mismo el uso del lenguaje orientado a la constatación de un determinado estado de cosas en el mundo objetivo que el uso del lenguaje orientado al establecimiento de un orden regulativo para las relaciones en conformidad con la validez de una norma. De ahí la importancia de la incorporación de la perspectiva de la tercera persona en el argumento moral, pues supone a un destinatario —individual o grupal— en tanto instancia activa y competente en la aprehensión del razonamiento y de quien se espera su aceptación. A diferencia de una perspectiva en la primera persona, el posicionamiento en la perspectiva de un agente externo supone que la información que se presenta en la comunicación del argumento es pública, que puede ser compartida, entendida y se pretende que sea aceptada. En ese sentido, lejos de cualquier «solipsismo moral y cognitivo» (Guariglia, 1996:48) para poder valorarse la objetividad del juicio moral se lo presenta de este modo.

Un punto más a propósito de la conclusión. Aunque parece claro, vale agregar que cuando la conclusión de un argumento moral tiene forma valorativa «x es bueno» allí queda implicada —*qua supra* argumento— «S debe realizar x» (Nunes y Salvadori, 2020); concluir una proposición evaluativa —por ej., «La eutanasia es mala»— contienen una equivalencia normativa directa en la proposición prescriptiva —por ej., «La eutanasia no debería permitirse/ aceptarse/ realizarse»—. El dar cuenta a través de razones que justifican una evaluación en la conclusión insta a un compromiso con un curso de acción.

Volviendo a la distinción entre niveles, podemos observar que cuando la validez de la norma se pone en duda, el nivel de análisis es distinto porque ello nos conduce a observar el criterio de corrección de la norma. Esto constituye un problema serio en sí mismo porque el mismo puede aludir a justificaciones de naturalezas muy distintas entre sí, por ej., deontológicas o consecuencialistas. El nivel superior inmediato al nivel moral denota el punto donde se proveen las garantías últimas en caso de que la norma fuera cuestionada, ese es el *nivel de los principios*. Allí pueden darse dos tipos de cuestionamientos,

uno en general y otro sobre el fundamento. Un relativista y un escéptico moral son representantes paradigmáticos del primer tipo en tanto podrían cuestionar la moralidad *in toto*, mas dicho cuestionamiento abre una serie de interesantes discusiones de metaética que aquí nos exceden y que nos alejarían del análisis de argumentos morales propuesto (Singer, 1995). Por otro lado, el cuestionamiento sobre los fundamentos sí es un punto, pues el mismo atiende a problemas vinculados con la garantía respaldatoria de las normas y esto está vinculado con las estructuras del razonamiento moral. Así visto, se puede trazar una cadena de razonamientos que nos retrotraiga a la discusión de principios llevándonos sin más a las teorías morales. Empero, la reflexión sobre principios ya no es argumentación moral sino teoría moral. Ahora bien, el hecho de que podamos retrotraernos debería estar lejos de conducirnos a cualquier posible escepticismo radical pues existe una importante cantidad de principios de morales fuertemente arraigados y aceptados en la mayor parte de las sociedades modernas.

Por último, siguiendo a Guariglia y Vidiella (2011:43), los razonamientos morales poseen una estructura análoga a razonamientos de otros campos en el sentido que retroceden de lo que se afirma en la conclusión a las razones que pretenden dar sustento. Ese camino de análisis retrospectivo se puede continuar hasta llegar a los principios de carácter universal que brindarían la garantía última. En analogía con el juicio teórico, el juicio moral puede ser visto como el *explanandum* para el que se requiere la norma universal — *explanans*— en la que está contenido (Guariglia y Vidiella, 2011:43). En tal sentido, como sostiene Romero, antes que responder a la pregunta monológica *¿qué debo hacer?* «la argumentación moral, por lo tanto, estaría orientada a dar razones que intenten responder a la pregunta de *¿por qué debo?»* (2020:172). De esta forma, la perspectiva monológica del razonamiento moral amplía sus fronteras cuando pasa de la dimensión personal a la interpersonal como espacio donde se traza la fundamentación de aquellas normas que regulan las relaciones interpersonales y que se comunican en un mismo contexto. Allí la argumentación puede ser vista como una práctica mediante la cual los participantes llegan a acuerdos intersubjetivamente válidos en los que no solo se pueden indagar cómo se establecen ciertas normas morales o jurídicas, sino que también se explicitan las razones y los principios que son dados en respaldo. Ahora bien, en el armado de una estructura de análisis argumentativo, todo esto bien podría quedar explicitado.

En relación con los puntos antes mencionados, se puede proveer un método de análisis argumentativo, necesario, si queremos hacernos de un análisis con algún grado importante de detalle sobre cómo se relacionan los distintos tipos de proposiciones. Para ello precisamos hacer de un esquema de estos

argumentos. Un esquema del argumento constituye la representación de un argumento en sus aspectos lógicos relevantes. En la representación se puede establecer una calificación a partir de preguntas que refuercen la relación lógica que vincula premisas con la conclusión.

## 1.

Podemos hacernos de un esquema básico de los argumentos morales como una representación de la forma en la que pueda observarse la interrelación que existe entre diferentes tipos de proposiciones y el tipo de inferencia que se realiza. Tomemos un primer ejemplo de argumento moral simple para su análisis.

### Ejemplo 1

*Cuando se autoriza la eutanasia para personas que atraviesan un prolongado coma profundo se autoriza la acción de propiciar unilateralmente la muerte de otro. De esto se sigue que la eutanasia no debe ser autorizada.*

1. Se autoriza la eutanasia ( $E$ )  $\rightarrow$  se autoriza matar a una persona ( $M$ )
2. No debe autorizarse ( $E$ )      Se sigue a partir de (1)

Estructura 1

1. ( $E$ )  $\rightarrow$  ( $M$ )
2. No debe autorizarse ( $E$ )      (1)

Muchas veces, los argumentos morales toman forma de entimemas donde queda implícita alguna/s premisa/s o la misma conclusión. Allí, el procedimiento de explicitación permite clarificarlo en sus premisas y en su formación. Dentro del esquema de un razonamiento moral podemos reconocer si el mismo está bien formado o no (Feldman, 1999:385 y ss). El argumento dado en el ejemplo anterior está mal formado. En el conjunto de premisas solo aparece explicitada ( $PE$ ) la premisa no moral (en adelante,  $P\rightarrow M$ ) y nada relaciona a la conclusión con  $P$ . Para hacernos de una forma bien formada precisamos incorporar la premisa moral ( $PM$ ) para que nos permita clarificar las relaciones que están dadas. Aquí, la premisa implícita (En adelante,  $PI$ ) es donde está dada la norma.

## Ejemplo 1b

1.  $(E) \rightarrow (M)$  (PE) (P $\neg$ M)
2. (M) *implica realizar una acción incorrecta* (I) (PI) (PM)
3. No debe autorizarse (E) (I) y (2)

## Estructura 1b

1.  $(E) \rightarrow (M)$
2.  $(M) \rightarrow (I)$
3. No debe autorizarse (E) (I) y (2)

La estructura dada en 1b, muestra una forma bien formada la cual está compuesta del siguiente modo:  $P_1, P_2 \vdash C$ . La premisa  $P_1$  incorpora el contenido no moral aportando de modo explícito aquello que puede ser verdadero o falso. En  $P_2$  se explicita un contenido moral que estaba implícito y que conecta a las premisas con la conclusión. La estructura dada en 1b muestra con más claridad la relación establecida y puede asociarse en la lógica clásica con la estructura del razonamiento válido de la transitividad, a saber:

$\varphi \rightarrow \psi, \psi \rightarrow \chi \vdash \varphi \rightarrow \chi$  Ahora bien, si la información provista en las premisas es razonable y es poco probable que nueva información implique la anulación de la conclusión, entonces, ello impacta en el grado de aceptabilidad del argumento. Empero, la relación entre el nivel descriptivo y el nivel prescriptivo tiene una cierta porosidad en la medida que entrecruza niveles donde en cada uno pueden establecerse ampliaciones semánticas a los términos que los componen.

Los esquemas tienen peso en sus conclusiones basados en razones que siempre pueden ser derrotadas a la luz de nueva información o evidencia. Por tanto, tales esquemas nunca serán absolutos o definitivos sino cerrados en algunos contextos en relación con la información disponible. (Nunes y Savadori, 2020:115)

Visto así, nueva información es posible que implique revocar la conclusión y deba deshacer la inferencia realizada.

La argumentación se da en contextos deliberativos donde las razones y justificaciones que proveen las premisas pueden ser abiertamente cuestionadas, especialmente si incorporan conceptos que a su vez precisan de una justificación interna con el propósito de que dicha premisa sea aceptada y se asegure externamente lo que puede ser un argumento válido o uno considerablemente convincente. Ambas premisas, las cuales responden a niveles diferentes pueden ser cuestionadas. Nunes y Savadori proponen complementar dialó-



gicamente la evaluación de los esquemas de razonamiento a partir de un conjunto estándar de cuestionamientos críticos: «(a) ¿La acción realmente tiene esa característica? (b) ¿Es esa característica relevante para la acción y/o contexto? (c) ¿Es la única característica relevante de la acción? (d) ¿Por qué es considerada moralmente buena/mala?» (2020:115).

A propósito de estos cuestionamientos, revisemos nuevamente nuestro argumento 1b. Teniendo en cuenta la pregunta (a) dirigida a  $p_1$ : ¿la acción de autorizar la eutanasia realmente tiene esa característica?; con respecto a (b) ¿es esa característica relevante para un contexto en el que existen pacientes en esa condición salud?; la pregunta (c) podemos dirigirla a  $p_2$ : ¿siempre es incorrecto matar o hay circunstancias que le quitan esa adjetivación moral?, ¿cuáles serían las circunstancias si las hubiere? El cuestionamiento provee un criterio de aceptabilidad de las partes que redundan en un esquema argumentativo mejor donde incluso podemos explicitar el principio. En ese sentido, Guariglia sostiene que:

cada uno de estos cuestionamientos desemboca en un entendimiento razonable a partir de datos disponibles, de modo que ambos interlocutores logren la suficiente certidumbre de que las afirmaciones de existencia de los hechos, las capacidades, las acciones, etc. responden razonablemente a la concepción compartida de la realizada física y social. (1996:49-50)

El cuestionamiento de la validez de la norma general nos lleva al nivel de los principios.

Estructura 1c

1.  $(E) \rightarrow (M)$  (PE) (P¬M)
2. (M) implica realizar una acción incorrecta (PI) (PM)
3. *Nunca ninguna persona debe darle muerte a otra* (PI) (PM – principio)
4. No debe autorizarse (E) (1) y (2)

La evaluación del argumento moral debería estar dada luego de intentar explicitar lo que fueren las premisas implícitas necesarias que refuerzan al argumento, dándole con ello mayor claridad a la relación lógica y moral que se pretende trazar —*contrario sensu* se puede estar propiciando la falacia del hombre de paja—. Si las premisas no morales son verdaderas, las morales son correctas, la forma es válida o convincente, se sigue como consecuencia la conclusión en un sentido fuerte:

Por cierto, poder establecer la verdad o falsedad del contenido narrativo que resume los hechos tiene una relevancia determinante para poder pasar al *segundo nivel* que es el *nivel moral* propiamente dicho. Por su estructura, el argumento moral permite extraer consecuencias del nivel epistémico, solamente si éste es verdadero. Solo entonces puedo sostener el juicio moral (...) Ahora bien, «tener relevancia» no quiere decir que la conclusión moral se siga del nivel epistémico. Para que la conclusión moral se pueda seguir debe haber una norma o una premisa moral que esté implícita y se explicita cuando el argumento lo requiera. (Guariglia y Vidiella, 2011:41)

En conclusión, dentro del esquema 1b y 1c del razonamiento moral podemos reconocer en el conjunto de premisas una justificación de la conclusión como juicio moral. Pasemos a un argumento nuevo.

### Ejemplo 2

*Dado que desde marzo 2020 el hecho de estar sin barbijo en espacios públicos cerrados es riesgoso tanto para uno como para los demás y teniéndose en cuenta que las acciones que ponen en riesgo a uno y a otros son malas, es evidente que es malo que alguien esté sin barbijo en lugares públicos cerrados.*

#### Estructura 2

1. Cualquiera que esté en un espacio público cerrado (P) sin barbijo ( $\neg B$ ) es riesgoso (R)
2. *Todo (R) es malo (M)*
3. Cualquiera (P) ( $\neg B$ ) es M / o Nadie (P) debería estar ( $\neg B$ )

#### Estructura 2b

1. La acción (P  $\wedge$   $\neg B$ ) tiene la propiedad R
2. *Toda acción con la propiedad R es moralmente M*
3. La acción (P  $\wedge$   $\neg B$ ) es moralmente M (1) y (2)

#### Estructura 2c (patrón)

1. La acción x tiene la propiedad R
2. *Todas acciones con la propiedad R es M*
3. La acción x es moralmente M (1) y (2)

#### Estructura 2d

1.  $\forall x (\forall y (Py \wedge Exy \wedge Bx) \rightarrow \forall z Rxz)$  Código
2.  $\forall x \forall y (Rxy \rightarrow Mx)$  Exy: x se ubica en y

3.  $\forall x (\forall y (Py \wedge Exy \wedge Bx) \rightarrow Mx)$       (1) y (2) Px: x es espacio público  
 3.1  $\forall x (\forall y (Py \wedge Exy) \rightarrow Bx)$       (3)      Bx: x tiene barbijo  
 rxy: x es un riesgo para y

En nuestro ejemplo, la conclusión inferida subordina el caso singular en cuestión a una norma general entendida como válida —dada en  $P_2$ —. Ahora bien, aquí tenemos un argumento bien formado y fuerte en el sentido que sus premisas son razonables. La valoración moral que aparece en la conclusión en el paso (3) puede reescribirse como una proposición prescriptiva en otras alternativas equivalentes tales como en (3.1) «Todos deben permanecer con barbijo en espacios públicos». En todos los casos se repite un patrón en la estructura de los argumentos morales simples (Feldman, 1999:386), a saber:

1. La acción x tiene la propiedad r
2. *Toda acción con la propiedad R está mal/ es incorrecta*
3. La acción x está mal/ es incorrecta      (1) y (2)

Más arriba habíamos llamado a nuestros ejemplos argumentos simples porque estaban compuestos de un modo estándar, por dos premisas y una conclusión. Un argumento moral complejo, como veremos, incorpora explícitamente dentro de sus premisas al principio moral. Pasemos a un último ejemplo de argumento moral complejo en un esquema consecuencialista.

### Ejemplo 3

Las únicas alternativas que tenían los gobiernos en 2020 eran hacer una aplicación masiva de vacunas contra el SARS-COV 2 aún en fase experimental —fase 3— o esperar mucho tiempo hasta que termine la investigación de los laboratorios. Si se optaba por la primera de las alternativas, el beneficio de las poblaciones era que si la vacuna tenía por encima de un 50 % de eficacia evitaban contraer la enfermedad de forma severa; no obstante, en la vacunación masiva habría un daño esperable por los efectos adversos que surgirían a causa de ser una investigación en proceso. No vacunar implicaba imprimir políticas de aislamiento, restricciones de circulación que impactaban en la economía y aceptar que se conviviría con una enfermedad muy contagiosa, sin tratamiento, con un índice de mortalidad del 5 % y una capacidad seria de generar saturaciones en la demanda de atención médica, camas y respiradores. El valor general de los daños que reporta la enfermedad es mayor y al que reportan los daños de la vacunación masiva. Era incorrecto no optar por una vacunación masiva. El principio de utilidad nos indica que debe procurarse el

aumento del beneficio o bien la disminución del daño. De allí que todos los gobiernos debían vacunar masivamente y llevar adelante políticas de promoción de la vacuna.

Estructura I

1. Las alternativas disponibles para G son vacunar (v) o esperar (E)
2. Los daños (D) de v son  $D_1, D_2 \dots D_n$
3. Los beneficios (B) de v son  $B_1, B_2 \dots B_n$
4. Los daños E son  $D'_1, D'_2 \dots D'_n$
5. Los beneficios de E son  $B'_1, B'_2 \dots B'_n$
6. El valor general de v es mejor que el valor general de E
7. *Principio de utilidad*
8. G debería optar por v (I) – (7)

Esquema Ib

1.  $v \vee E$
2.  $v \rightarrow D_1, D_2 \dots D_n$
3.  $v \rightarrow B_1, B_2 \dots B_n$
4.  $E \rightarrow D'_1, D'_2 \dots D'_n$
5.  $E \rightarrow B'_1, B'_2 \dots B'_n$
6.  $(B_1, B_2 \dots B_n > B'_1, B'_2 \dots B'_n) \wedge (D_1, D_2 \dots D_n < D'_1, D'_2 \dots D'_n)$
7. *Principio de utilidad*
8. G debería optar por v (I) – (7)

Estructura Ic Simplificación del esquema consecuencialista en un argumento simple sin el principio (7) como premisa.

1. Si se realiza la acción v, entonces se produce/n la/s consecuencia/s c/  $C_1, C_2 \dots C_n$
2. c es moralmente mejores/deseables/preferibles  
 $C_1, C_2 \dots C_n$  son moralmente mejores/deseables/preferibles
3. La acción v es moralmente mejor/deseable/preferible (I) y (2)

Esquema Id

1.  $v \rightarrow c / C_1, C_2 \dots C_n$
- 1.1  $\neg v \rightarrow c' / C'_1, C'_2 \dots C'_n$
2.  $C_1, C_2 \dots C_n$  son moralmente mejores/deseables/preferibles
- 2.2.  $C'_1, C'_2 \dots C'_n$  no son moralmente mejores/deseables/preferibles
3. La acción v es moralmente buena/deseable/correcta (I) – (2.2)

## A modo de cierre

No hay un único modo para traducir y reconstruir los argumentos morales. A la hora de hacerlo, podemos utilizar las herramientas conceptuales que nos brindan la lógica o alguna teoría de la argumentación. En el presente texto nos hemos concentrado en proponer una descripción de las partes y los niveles que se ponen en juego cuando se da un argumento moral. Asimismo, esbozamos cómo la reconstrucción de los mismos constituye un procedimiento importante para la identificación de premisas y conclusión, clarificación a través de la adición de premisas implícitas y la visualización de su estructura para advertir y evaluar sus debilidades o fortalezas. Si consideramos la vinculación que tienen los argumentos de este tipo con nuestras creencias morales, entonces, podemos advertir de la importancia que tiene la tarea de análisis aplicada a ellos.

## Referencias bibliográficas

- Abal, F.G. (2020). Guía: Introducción a la argumentación moral. Sedici.
- Ausín, T. y Peña, L. (2011). Premisa/Conclusión. En Vega Reñon, L. y Olmos Gómez, P., Compendio de lógica, argumentación y retórica (pp. 479-481). Trotta.
- Busdygan, D. (2017). Construcción política en el desacuerdo sobre el aborto. *Descentrada*, 1(2), 17-29.
- Copi, I. (2009). Introducción a la Lógica. Eudeba.
- Feldman, R. (2013). Reason and Argument: Pearson New International Edition. Pearson Higher.
- Gamut, L.T.F. (2009). Lógica, Lenguaje y Significado: Introducción a la Lógica. Eudeba.
- Guariglia, O. (1996). Moralidad. Ética universalista y sujeto moral. Fondo de Cultura Económica.
- Guariglia, O. y Vidiella, G. (2011). Breviario de ética. Edhasa.
- Legris, J. (2001). Deducción y representación. Economizarte.
- Moretti, A. (2020). En sayos analíticos. Analizar, enunciar y deducir. SADAF.
- Nunes, E. y Salvadori, M. (2020). Some characteristics and complexities of moral argumentation. *Prometheus-Journal of Philosophy*, (32), 109-126.
- Plantin, C. y Muñoz, N. (2011). El hacer argumentativo. Biblos.
- Romero, J. (2020). El desarrollo dialógico de la argumentación moral y sus relaciones con la ética del discurso. *Oximora. Revista internacional de Ética y Política*, (16), 158-180.
- Singer, P. (Ed.). (1995). Compendio de ética. Alianza.
- Strawson, P.F. (2011). Introduction to logical theory. Routledge.
- Vega Reñon, L. y Olmos Gómez, P. (2011). Compendio de lógica, argumentación y retórica. Trotta.

## 11. Experimentos mentales y el rol de la historia

*Santiago Truccone-Borgogno*

De modo similar al que en las ciencias químicas, físicas y naturales se realizan experimentos en laboratorios con el fin de aislar elementos y analizar diferentes reacciones, los filósofos prácticos construyen diferentes clases de experimentos mentales o ejemplos con el objetivo de aislar propiedades relevantes que hacen surgir o permiten testear principios normativos (morales y políticos) y demás juicios morales considerados.

Por tanto, el uso de experimentos mentales, en particular de ejemplos hipotéticos, es una herramienta valiosa en la investigación filosófica, dado que nos permite lidiar exclusivamente con conceptos y posibilidades sin necesidad de dar cuenta de hechos concretos, probabilidades y de la posibilidad de que las cosas sean empíricamente problemáticas. Es decir, los experimentos mentales nos permiten ignorar las contingencias (cf. Shue, 2009:307). Quizás sea por estas razones que su uso es una práctica generalizada en la filosofía política y moral contemporánea.

En este trabajo, sin embargo, argumentaré que el uso exclusivo de experimentos mentales es problemático al momento de lidiar con algunos problemas importantes de filosofía práctica. No me estoy refiriendo a problemas teóricos específicos tales como el problema de la no identidad (Parfit, 1984 – cap. 16), sino a cuestiones generales como, por ejemplo, la referida a la orientación temporal de la justicia. ¿Deben las injusticias históricas influir en lo que debemos hacer o en cómo nuestras instituciones deben ser diseñadas? O, en cambio, ¿son las necesidades presentes de las personas actualmente existentes las que deben determinar lo que la justicia requiere?<sup>1</sup>

Las injusticias históricas y sus efectos duraderos dan lugar a circunstancias en las que los intereses de diferentes personas y grupos entran en conflicto. Dichos conflictos de intereses, por tanto, justifican la tarea de buscar un principio de justicia que ayude en su solución dado que el propósito de los principios de justicia es, precisamente, determinar cómo las personas deben actuar (o vivir) y cómo deben diseñarse las instituciones públicas en situaciones de conflicto de intereses.

---

1 Para un análisis detallado de distintos problemas relacionados con injusticias históricas y sus posibles efectos en el presente y el futuro véase Meyer (2001, 2004, 2006, 2020); Truccone-Borgogno (2021a, 20221b, 2022a, 2022b, 2022c).

Al igual que con relación a otros problemas en filosofía práctica, aquí nos encontramos con un asunto en el que nuestros juicios morales considerados entran en conflicto. Por un lado, existe una fuerte intuición que indica que los principios de justicia deben priorizar las necesidades presentes de las personas presentes. Por otro lado, también tenemos una fuerte intuición que afirma que aquellos que son injustamente dañados merecen y deben recibir alguna forma de compensación. ¿Cómo debemos proceder ante una situación de este tipo?

Una estrategia habitual para intentar responder a esta pregunta es apelar a ejemplos hipotéticos, como mostraré más abajo. Apelar a ejemplos hipotéticos en asuntos como este intenta colocar al juez (moral) en una situación donde sus propios intereses no están en juego y, de ese modo, se crea una especie de escenario donde se garantizan condiciones de imparcialidad. Sin embargo, argumentaré en este trabajo que el uso de dichos ejemplos es problemático dado que, por su mismo diseño, excluye la posibilidad de que la historia pueda tener algún valor en lo que la justicia requiere en el presente. Por tanto, en lugar de servir como herramienta donde las consideraciones a las que se lleguen puedan catalogarse como imparciales entre concepciones en pugna, en realidad, favorecen un tipo específico de respuesta. De ese modo, no dan cuenta del propósito de servir como dispositivos que garanticen imparcialidad en la formulación de juicios morales.

Para apoyar mi argumento, en la sección I, mostraré qué lugar ocupan los experimentos mentales en las investigaciones filosóficas prácticas que se conducen a través del uso del método de equilibrio reflexivo. Una vez que haya mostrado que los experimentos mentales pueden cumplir la función de servir como dispositivos o test formales de validez de los juicios normativos, en la sección II, me adentraré en el problema del rol de la historia en nuestras discusiones sobre justicia contemporánea. Mostraré cómo la apelación a experimentos mentales favorece cierto tipo de respuestas en las que la historia no puede tener lugar relevante. En la sección III muestro que, si uno se apoya en historias reales, nuestros juicios morales considerados con relación al rol de la historia en discusiones de justicia parecen diferir de aquellos que obtenemos al usar ejemplos hipotéticos. Concluyo el ensayo afirmando que, por tanto, para ciertos asuntos, el uso de experimentos mentales (hipotéticos) debe complementarse necesariamente con el de ejemplos ostensivos o reales. De otro modo, no habremos alcanzado un equilibrio reflexivo capaz de justificar nuestros principios morales.

## 1. Equilibrio reflexivo y experimentos mentales

Un modo de proceder ante el tipo de conflictos reseñado en la sección anterior, quizás el más extendido en la filosofía analítica práctica contemporánea, es a través del uso del método llamado *equilibrio reflexivo*. Este método se denomina *equilibrio* porque al final del proceso nuestros principios y demás juicios morales considerados coinciden. Asimismo, se llama *reflexivo* porque conocemos tanto qué principios se ajustan a nuestros juicios morales considerados (intuiciones) como también las premisas de su derivación (cf. Rawls, 1971:20). Según la reconstrucción de Norman Daniels el equilibrio reflexivo se alcanza cuando existe coherencia entre (1) un conjunto de principios normativos, (2) un conjunto de juicios morales considerados (intuiciones), y (3) un conjunto de teorías de fondo relevantes (Daniels, 2016 – Sección 1).<sup>2</sup>

Sin embargo, es plausible sostener que también existe un cuarto elemento que debe ser incorporado para alcanzar equilibrio reflexivo de manera que las teorías y principios morales puedan estar justificados. Me refiero a lo que Carlos Nino llama (4) *test formal de validez* (Nino, 1991:69-71). En una interpretación, *la posición original* de Rawls también puede ser entendida como un test formal de validez. La posición original se refiere a un conjunto de circunstancias diseñadas para adoptar principios de justicia desde un punto de vista imparcial que, a su vez, estén basados en el autointerés de nuestros representantes (Rawls, 1971 – Cap. 3). Nino afirma que, si el diseño de la posición original permitiese que estén justificadas algunas prácticas que son ampliamente rechazadas en el presente, tales como la esclavitud, tiene que suponerse que el dispositivo pasó por alto algún aspecto importante sobre cómo los juicios morales deben ser formulados (Nino, 1991:70). En este caso, dado que el dispositivo permite conclusiones implausibles tenemos razones para revisarlo.

Los tests formales de validez deben tratar de alcanzar dos objetivos. Primero, su diseño debe evitar sesgos al proferir juicios morales. Deben disminuir tanto como sea posible la influencia de factores que distorsionen el juicio al formular consideraciones morales, tales como el uso de la fuerza, la autoridad, o la posibilidad de ser castigados o premiados en función de cómo se evalúa la situación a juzgar. Segundo, deben ser capaces de generar o identificar principios morales (o al menos consideraciones relevantes para su formulación) que ayuden en la solución de problemas de la vida social.

---

<sup>2</sup> Esta caracterización corresponde con lo que se llama equilibrio reflexivo *amplio*. Este difiere del llamado equilibrio reflexivo *estrecho* dado que este último solo requiere alcanzar coherencia entre (1) principios e (2) intuiciones (cf. Daniels, 2016 – Sección 3; Rawls, 1971:46-53).



La posición original de Rawls no es, por supuesto, el único dispositivo al que pueda acudirse. Hablar de observadores imparciales, virtuosos o ideales, ponernos en los zapatos de otros, imaginarnos a nosotros mismos en una situación de deliberación pública, la presunción de igualdad (básica) entre las personas, entre otros, son todos intentos de encontrar un test formal de validez de los juicios morales. A esta categoría pertenecen también los experimentos mentales. Un experimento mental, como afirman Kimberley Brownlee y Zofia Stemplowska, «es un proceso dividido en etapas que involucra (1) una visualización mental de algún escenario específico con el propósito de (2) responder a otras preguntas más generales y, al menos, parcialmente independientes de los estados mentales» (Brownlee y Stemplowska, 2017:25).<sup>3</sup>

Estas autoras afirman que un experimento mental bien formulado debe cumplir con cuatro condiciones: (i) validez, (ii) ausencia de petición de principio, (iii) encuadre narrativo moderado, y (iv) inclusión de alternativas rudimentarias. El requisito de (i) *validez* refiere al hecho de que los experimentos mentales tienen que ser susceptibles de proporcionar alguno de los siguientes resultados: ellos pueden elucidar qué juicio es moralmente requerido en el caso concreto, o —adicionalmente— ellos pueden proporcionar una hipótesis sobre las razones por las que ese juicio es correcto. Solo en el segundo caso, «el experimento mental puede “ser traducido” en un argumento» (Brownlee y Stemplowska, 2017:31). El requisito de (ii) *ausencia de petición de principio* demanda que los «experimentos mentales no asuman la respuesta a la pregunta que se formula con ellos» (31). El requisito de (iii) *encuadre narrativo moderado* tiende a evitar que la misma narrativa favorezca o desaliente alguna respuesta específica. Finalmente, el requisito de (iv) *inclusión de alternativas rudimentarias* es importante sobre todo cuando se opera con ejemplos imaginarios. Los ejemplos *imaginarios* «involucran posibilidades lógicas que solo pueden tener lugar en un mundo diferente del nuestro» (Jamieson, 1993:684). Este tipo de experimentos mentales se caracterizan por alterar algunas leyes sobre cómo el mundo, de hecho, funciona. Por lo tanto, cuando se diseña un experimento mental de este tipo es necesario proporcionar una alternativa a la característica del mundo real que no está presente en nuestro ejemplo. Esto es relevante dado que, si no proporcionamos una alternativa, podríamos continuar asumiendo que la característica o propiedad inexistente aún está presente en el ejemplo (cf. Brownlee y Stemplowska, 2017:30–37).

Los experimentos mentales, por tanto, cuando son pensados como dispositivos que sirvan de tests formales de validez de los juicios morales deben evitar que su diseño mismo o estructura favorezca una de dos o más posibles

---

3 En este capítulo, todas las traducciones del inglés al castellano son propias.

consideraciones plausibles en conflicto. Deben garantizar, en otros términos, imparcialidad al momento de proferir juicios morales considerados. Asimismo, deben modificarse cuando se identifica que su diseño o bien permite respuestas implausibles o bien favorece un tipo de respuesta sobre otra plausible. Los experimentos mentales también deben estar en equilibrio reflexivo con los principios normativos, intuiciones y teorías de fondo.

## 2. Experimentos mentales y el valor moral de la historia

En la sección anterior he afirmado que los experimentos mentales pueden cumplir la función de servir como dispositivos o test formales de validez de los juicios normativos. En esta sección incursionaré en el problema del valor moral de la historia para la justicia.<sup>4</sup> Mostraré cómo la apelación a experimentos mentales favorece cierto tipo de respuestas en las que la historia, en particular las injusticias pasadas, no tiene valor moral en el presente.

¿Pueden las injusticias pasadas afectar aquello que en el presente es requerido por consideraciones de justicia? Veamos el siguiente experimento mental proporcionado por David Lyons:

a) Supóngase que somos ocupantes de una isla remota. Nos las hemos arreglado para usar la tierra y todos sus otros recursos entre nosotros, y vivimos de manera confortable, con algunos bienes menos perecederos que reservamos para las temporadas de sequía. Un día, unos naufragos son arrastrados a nuestras costas luego de que su buque se destruyera. Son huéspedes que no hemos invitado, pero, también, involuntarios. No hay perspectivas de que se retiren de forma segura y no tienen recursos más allá de su capacidad para trabajar. Sin embargo, ellos no son agresivos. (Lyons, 1981:369)

¿Qué debe hacerse en esta situación? Por supuesto, no podemos echar a estos naufragos hacia el mar. No obstante, su llegada crea un problema en la distribución de recursos que requiere solución. Si las cosas se mantienen sin cambios, como nota Lyons, los recién llegados terminarán condenados a vivir

---

4 No me refiero ni al problema sobre cómo evaluar si cierto hecho o conjunto de hechos pasado pueden ser catalogados como injustos, ni al asunto sobre cómo elegir de manera justificada entre distintas interpretaciones disponibles referidas al pasado. He analizado este problema en Truccone-Borgogno (2021a – Cap. 1) La pregunta que me interesa analizar en este capítulo es previa a esos asuntos. Me interesa indagar sobre si algo que sucedió en el pasado (independientemente de qué interpretación concreta se le dé) puede tener influencia en lo que debemos hacer o en cómo las instituciones deben ser diagramadas actualmente.

en una posición subordinada tanto económica como socialmente (1981:380). Para evitar tal consecuencia, parece que medidas de redistribución son requeridas. En este mundo sin injusticia pasada, consideraciones de justicia distributiva hablan en favor de otorgar una porción equitativa de tierra a los recién llegados. Esta historia sugiere que la justicia debe tener una orientación prospectiva (*forward-looking*). Como conceptualiza Leif Wenar, «las justificaciones prospectivas, (...), miran sólo a las razones que tenemos para hacer de ahora en más al mundo mejor de lo que es» (2006:297).

Consideremos ahora una segunda historia:

b) Supóngase que los naufragos que llegan a nuestras costas no son amistosos y cooperativos sino agresivos y dominadores. Nosotros tratamos de ser hospitalarios, pero ellos no corresponden. Nos engañan, matan a muchos de nosotros y fuerzan a los sobrevivientes a residir en una pequeña área de la isla lejos de nuestros hogares, mientras ellos se apropian de una parte desproporcionadamente grande de la isla, incluyendo a los sectores más deseables. (Lyons, 1981:373)

En esta segunda historia, los invasores agresivos deberían reparar aquellas injusticias que perpetraron en contra de los ocupantes de la isla. Aquí, luego de las injusticias, parece que hay razones para hacer que el estado de cosas sea como era antes de que los invasores lleguen. A diferencia de la primera historia, el segundo escenario sugiere que la justicia debería tener una orientación retrospectiva (*backward-looking*). Para Wenar, «las justificaciones retrospectivas están basadas en razones básicas que (supuestamente) tenemos que hacer que el presente coincida con el pasado» (2006:296).

La razón por la que, en el primer escenario, la justicia tiene una orientación prospectiva, mientras que, en el segundo, una mirando hacia el pasado, es que distintos aspectos de la justicia están implicados en cada uno de ellos. En el primero, solo aparecen preocupaciones de justicia distributiva mientras que, en el segundo, nos confrontamos también con consideraciones de justicia compensatoria. Sin embargo, aunque parece que siempre tenemos razones «para hacer de ahora en más al mundo mejor de lo que es», tal y como recomiendan los principios normativos fundados en justificaciones prospectivas, no es claro si, una vez que una injusticia fue cometida, siempre (y para siempre), tenemos razones «para hacer que el presente coincida con el pasado» —como los principios normativos justificados en consideraciones retrospectivas recomiendan.

Imaginemos, para seguir con la narrativa de Lyons, que la segunda historia (b) continúa como sigue: «Los descendientes de los piratas invasores comienzan a vivir armónicamente con nuestros descendientes, de manera que nadie

es privado de una parte equitativa de los recursos de la isla» (Lyons, 1981:374). En esta continuación de la segunda historia, Lyons piensa que nuestros descendientes no tienen reclamos adicionales por compensación en contra de los descendientes de los invasores. De acuerdo a esta posición, todas las injusticias (*wrongs*) que sufrimos no pueden ser reparadas luego de que morimos (Lyons, 1981:374). Por lo tanto, parece que las consideraciones retrospectivas dejan de tener peso moral en la historia en que, por un lado, las víctimas y perpetradores pasados murieron y, por el otro lado, las personas presentes viven en armonía en condiciones en las que todos tienen su parte equitativa de los recursos de la isla. De hecho, proporcionar compensación en tales circunstancias parece requerir una alteración a la distribución de recursos perfecta que, ahora, existe en la isla; hará que la isla sea distributivamente menos justa. Por esta razón, algunos filósofos afirman que las reparaciones por injusticias pasadas no deben permitirse en aquellas situaciones en las que proporcionar compensación es incompatible con preocupaciones prospectivas de justicia distributiva (cf. Wenar, 2006:402; Vernon, 2003:544).

Este es el tipo de consideraciones que Jeremy Waldron parece apoyar en favor de su tesis de la superación (*supersession*) de las injusticias históricas. Al comentar la segunda historia de los invasores afirma:

Lo que se requiere es una distribución moralmente apropiada a las circunstancias *presentes*, a los recursos *presentes*, y a la población *presente* que no tiene otra opción que vivir en estos territorios; y no a un conjunto de circunstancias pasadas en las que todo era diferente. En este sentido, una injusticia histórica puede ser superada, y la reparación derrotada (*trumped*), por decirlo de algún modo, por principios de justicia directamente aplicados a las circunstancias presentes. (Waldron, 1992a:165 – énfasis en el original)

El punto de la discusión es que las razones retrospectivas para proporcionar compensación importan hasta que ellas entran en conflicto con razones prospectivas de justicia distributiva. Waldron fortalece su caso en favor de la superación de las injusticias pasadas empleando otro experimento mental en el que, en el pasado, con circunstancias de abundancia de recursos, un grupo *N* invade y se apropia de manera injusta de un pozo de agua que pertenece a otro grupo *S* sin reciprocidad. Esto es una injusticia que, por supuesto, requiere reparación. En aquel tiempo,  $\tau_1$ , el pozo de agua debió ser devuelto a sus dueños originarios. De repente, en  $\tau_2$ , hay un desastre natural que hace que todos los pozos de agua se sequen salvo el pozo que está ya siendo compartido por los dos grupos *N* y *S* (Waldron, 1992b:24–25). De acuerdo con Waldron, en tal caso, la injusticia fue superada por el cambio en las circunstancias

dado que, ahora, devolver el pozo de agua al grupo s condenaría a los miembros del grupo N a morir o, al menos, a vivir en condiciones muy deplorables (1992b:24-25). La plausibilidad de la tesis de la superación es fortalecida en esta versión del caso porque ahora, esto es, en el *presente*, ambos grupos están en circunstancias de necesidad extrema causadas por un desastre natural.

Pensemos nuevamente en el ejemplo de los invasores de nuestra isla. Si estábamos dispuestos a pensar que las razones retrospectivas dejaban de tener peso moral cuando, en el presente, cada parte tiene su porción equitativa de la isla, parece que tal conclusión es más fuerte en el ejemplo hipotético en el que los descendientes tanto de los invasores como de los ocupantes originarios de la isla están, debido a un desastre natural, en condiciones de necesidad extrema. Entendida de este modo, la tesis de la superación da lugar a un principio general de acuerdo con el cual la orientación temporal de la justicia cambia cuando hay situaciones de necesidad. En particular, las razones retrospectivas de justicia compensatoria dejan de tener peso tan pronto como entran en conflicto con razones prospectivas de justicia distributiva basadas en la necesidad.

### 3. Historia y justicia

En la sección anterior argumenté que el principio normativo al que hemos llegado parece estar en equilibrio reflexivo con los juicios morales considerados concretos (intuiciones) que se obtienen al pensar sobre hechos involucrados en los experimentos mentales ideados por Lyons y Waldron. Por tanto, tal principio parece estar justificado: ante situaciones de necesidad (y, tal vez, también otras situaciones de justicia distributiva) la historia no tiene peso moral. Ahora bien, ¿qué sucedería si, en lugar de descansar en experimentos mentales del tipo ideado por Lyons y Waldron, analizamos casos reales?, ¿estamos seguros de que nuestras intuiciones serían similares? En esta sección intentaré mostrar que nuestros juicios morales considerados en relación al rol de la historia en discusiones de justicia parecen diferir de aquellos que obtenemos al usar ejemplos hipotéticos.

Al igual que la *posición original* de Rawls (1971), tanto los ejemplos de los *náufragos* e *invasores* de Lyons (1981) como los de los *pozos de agua* de Waldron (1992b) son experimentos mentales que tienen el objetivo de servir como *tests formales de validez* de nuestros juicios morales. En particular, los escenarios delineados por Lyons y Waldron son ejemplos hipotéticos. Esta clase de ejemplos «representan situaciones que pueden haber tenido lugar o podrían tener lugar sin forzarnos a reescribir la física o a cambiar nuestras concepciones básicas sobre cómo el mundo funciona de hecho» (Jamieson,

1993:484). La ventaja de apoyarnos en esta clase de ejemplos es que, sin ser ejemplos imaginarios, permiten aislarnos de las circunstancias reales y llegar —en la mayoría de los casos— a soluciones con un mayor grado de imparcialidad. Usados como dispositivos de validez formal, ellos son una herramienta útil para evitar sesgos en la provisión de juicios morales.

Sin embargo, el uso de experimentos mentales, en particular el uso de los ejemplos hipotéticos es peligroso a los fines de evaluar el valor moral de la historia. Al ser generalizaciones abstractas, los ejemplos hipotéticos parecen no permitir una reflexión crítica sobre el rol de la historia y sobre cómo ella puede afectar a posibles soluciones de conflictos de intereses. En cierto sentido, este tipo de experimentos mentales parecen incurrir en una suerte de *petición de principio*, dado que pretende evaluar el valor moral de la historia de un modo que, de hecho, la excluye de un posible análisis (ver Ivison, 2000:370–372; ver también Waligore, 2016, 2018).<sup>5</sup>

Para ver este punto, permítaseme contar una breve historia, ahora, real:

En noviembre de 1878, el coronel Conrado Villegas captura al cacique indígena Pincén junto con varios miembros de su familia en las cercanías de Trenque Lauquen (provincia de Buenos Aires), y son apresados en la isla Marín García. Como resultado de ésta y de las campañas militares subsiguientes la mayoría de los miembros sobrevivientes de la comunidad del «viejo» cacique Pincén fueron forzados a emigrar al oeste y sur del territorio que actualmente pertenece a la Argentina (cf. Nagy, 2014:100). Aunque la historia de la comunidad es difícil de seguir, hay registro de que la nieta del «viejo» cacique Pincén, Marcelina Pincén y su pareja, Honorio Cejas, organizaron varias reuniones con los descendientes de los miembros de la comunidad del «viejo» cacique Pincén a finales de la década del 60. Sin embargo, los intentos de obtener el reconocimiento legal de la comunidad fueron interrumpidos debido a la política represiva iniciada por el Ministro de Bienestar Social del gobierno peronista José López Rega; política represiva intensificada durante la última dictadura militar de 1976–1983, al punto que, incluso, algunos miembros de la comunidad terminaron presos (cf. Nagy, 2014:103). Con el retorno de la democracia a la Argentina, Lorenzo Cejas–Pincén, bisnieto del «viejo» cacique Pincén, promovió el reconocimiento de lo que hoy se conoce como *Comunidad Indígena Cacique Pincén*.<sup>6</sup> Esta comunidad tiene como uno de sus objetivos recuperar las tierras de las que fueron despojados a finales

---

5 Esto es contrario el tercer requisito (iii) para formular experimentos mentales de manera correcta, como he reseñado en la sección I de este trabajo.

6 Tal comunidad obtuvo su reconocimiento legal en el Registro Nacional de Comunidades Indígenas el 14 de abril de 2015. Para una descripción detallada de la historia de esta comunidad, véase Nagy (2012). Ver también <http://datos.jus.gob.ar/dataset/listado-de-comunidades-indigenas/archivo/f9b57566-3e7c-4449-b984-49a26897eb77>.

del siglo XIX.<sup>7</sup> Aun así, varios miembros de la comunidad viven en condiciones precarias sin posibilidad de ver sus necesidades básicas satisfechas. Aún más, la comunidad nunca ha recibido los títulos legales de tierras con la calidad suficiente para que puedan vivir de su producto. (cf. Nagy, 2014:102-123)

¿Está justificada la pretensión de recuperar sus tierras en el presente? ¿Existen razones para compensar a comunidades indígenas como las de Pincén por la injusticia sufrida en el pasado? O, en cambio, ¿sus reclamos por recibir compensación han sido superados (*superseded*)?

A diferencia de los experimentos mentales proporcionados por Waldron y Lyons, la historia de Pincén, constituye un ejemplo *real* u *ostensivo*. Este tipo de ejemplos indican hechos o personas reales cuyas características son moralmente significativas (O'Neill, 1986:7). Una de las ventajas de utilizar este tipo de ejemplos como dispositivos para la obtención de juicios morales considerados radica en que presentan una situación concreta que, sin dudas, tiene relevancia práctica y debe ser resuelta por algún principio moral.

Al ser confrontado con este ejemplo, al menos en mi caso, la intuición indica que existen razones retrospectivas para que la comunidad de Pincén reciba alguna forma de compensación. La sucesión de injusticias sufridas por Pincén y su comunidad parece requerir reparación. Sin embargo, el problema es que, al igual que en el experimento mental de los pozos de agua de Waldron (o la historia que nos cuenta Lyons en la versión en que un desastre natural deja a todos los habitantes con pocos recursos), las circunstancias actuales difieren de aquellas presentes a finales del siglo XIX. Hoy en día, la población de Argentina es estimada en 46 200 000 de habitantes,<sup>8</sup> y no 4 044 911 como había en 1885 —año en que culminó la llamada Conquista del Desierto—.<sup>9</sup> En 2019, su densidad poblacional era de 14,6 habitantes por km<sup>2</sup>, y no 1,40 hab/km<sup>2</sup>.<sup>10</sup> Además, el país cuenta con más del 37 % de su población viviendo en condiciones de pobreza.<sup>11</sup> Estas condiciones no tienen perspectiva de mejora; todo lo contrario. El nuevo coronavirus (SARS-COV-2), usualmente llamado COVID-19, forzó al gobierno de Argentina a tomar medidas drásticas<sup>12</sup> que han conducido no

---

7 Artículo 3, sección B del Estatuto de la comunidad de 1995 (cf. Nagy, 2014:108).

8 <https://www.indec.gob.ar/>

9 Estos números surgen del censo de 1895. En ese momento había 90 000 personas indígenas.

30 000 de las cuales estaban «más allá del imperio de la civilización» (De la Fuente, 1898:XIX).

10 La densidad poblacional actual de la Provincia de Buenos Aires (donde la comunidad de Pincén está localizada) es 50 hab/km<sup>2</sup>.

<https://www.ign.gob.ar/NuestrasActividades/Geografia/DatosArgentina/DivisionPolitica>

11 <https://www.indec.gob.ar/indec/web/Nivel3-Tema-4-46>

12 Hasta, al menos el 10 de septiembre de 2020, con algunas pequeñas modificaciones, la cuaren-

solo a un decrecimiento en su economía sino también a un aumento en las condiciones de desigualdad social (cf. Bonfiglio, Salvia y Vera, 2020).

A la luz de estas nuevas circunstancias no es extraño que pueda afirmarse que, ahora, los habitantes de Argentina se encuentran viviendo en una situación de necesidad que hace que sea justo que las comunidades indígenas y los no-indígenas tengan que compartir aquellas tierras que pertenecían de manera exclusiva a los pueblos indígenas en el pasado y que los últimos ya no deban ser compensados. Si aceptamos el principio que emerge de haber considerado tanto el ejemplo de los pozos de agua de Waldron como el de los invasores de la isla de Lyons, las razones retrospectivas para compensar a comunidades como las del Cacique Pincén ya no tienen un fuerte peso moral. Por tanto, la injusticia pasada parece haber sido superada.

Mi propia reacción ante este caso difiere de aquella que me han provocado los experimentos mentales de Waldron y Lyons. A diferencia de aquellos casos hipotéticos, al ser confrontado con este caso real, no me resulta obvio que la injusticia pasada haya sido superada. De hecho, mi intuición indica que las razones para proporcionar compensación todavía tienen bastante peso moral en este caso. ¿Cómo es esto posible? Una posibilidad es que exista un desequilibrio en mi persona que hace que, ante casos con propiedades morales semejantes, mis juicios morales difieran. Otra posibilidad es que los ejemplos no compartan propiedades morales relevantes. Sin embargo, existe una tercera opción. Es posible que los ejemplos reales u ostensivos puedan hacer surgir juicios morales considerados, en particular, intuiciones, que no podrían surgir de ningún modo de ejemplos hipotéticos o imaginarios.<sup>13</sup>

No quisiera exagerar las conclusiones que estoy sugiriendo. No estoy afirmando que las investigaciones en filosofía práctica que descansan en experimentos mentales no dan cuenta de manera adecuada de lo que Nino ha llamado *test formal de validez*. Todo lo contrario, por las razones esgrimidas en las secciones anteriores, los ejemplos hipotéticos e imaginarios son de mucha importancia para elaborar y justificar —en muchas ocasiones— principios morales.<sup>14</sup> Sin embargo, en algunos asuntos concretos, los experimentos

---

tena general estaba todavía vigente. Una lista detallada de las medidas tomadas por el gobierno de Argentina para lidiar con los efectos en el país de la pandemia global, ver: <https://www.argentina.gob.ar/coronavirus/medidas-gobierno>.

13 Es muy probable que algunas personas no compartan el juicio moral considerado referido a que, en esta historia real, se debe compensación en este ejemplo real. Este hecho, sin embargo, no muestra que mi argumento es equivocado. No intento mostrar que los ejemplos reales muestran que la historia tenga peso moral en lo que debemos hacer actualmente, sino solo que —a diferencia de los hipotéticos— este tipo de ejemplos habilitan dicha posibilidad.

14 En otro lugar he argumentado con más detalle el valor de utilizar ejemplos hipotéticos e imaginarios en investigaciones morales (Truccone-Borgogno, 2018a). De hecho, he descansado en



mentales no pueden proporcionar todos los elementos necesarios para que un principio normativo esté justificado. Como afirma Nino, la práctica de formular juicios morales está destinada a ayudar a los individuos y grupos a disminuir sus conflictos y facilitar la cooperación social (Nino, 1991:67). Si el problema que hace que las personas y grupos puedan tener intereses en conflicto se debe a que no es claro el valor moral de la historia, no es posible que la respuesta adecuada a tal problema surja o sea disparada a través de un dispositivo —ejemplo hipotético— que no permite analizar si la historia puede o no tener un valor moral en lo que debe hacerse.

Con esto no quiero tampoco afirmar que no debamos utilizar experimentos mentales para analizar el valor moral de la historia en nuestras discusiones de justicia distributiva. En todo caso, el punto es que descansar en el uso de experimentos mentales no es suficiente. Si como test formal de validez de juicios morales vamos a utilizar ejemplos, en la medida en que queramos evaluar el valor moral de la historia, el uso de tales experimentos debe completarse con el de ejemplos ostensivos o reales. Para estos fines, el uso de ejemplos reales es inevitable para darle mayor fuerza justificatoria al principio normativo que estemos buscando. De otro modo, no será posible identificar algunas intuiciones y juicios morales considerados que, luego de que emerjan, también deben ser equilibrados con los principios normativos y teorías de fondo.

Ahora bien, si es verdadero que los ejemplos ostensivos pueden dar cuenta de ciertos juicios morales considerados que los experimentos mentales no, ¿por qué no descansar directa y completamente en su uso? La respuesta a esta pregunta debe comenzar por notar que, al igual que los experimentos mentales hipotéticos, los ejemplos reales también tienen debilidades. Primero, muy raramente ellos contienen todas las propiedades normativamente relevantes. Segundo, incluso si las contienen, «la relación entre el caso real y el problema moral que se esté analizando debe ser guiada por cierto entendimiento (necesariamente) indeterminado de los aspectos morales del ejemplo» (O'Neill, 1986:8). Tercero, dado que son ejemplos reales, es posible que uno pueda sentirse involucrado en ellos de un modo contrario a lo que la imparcialidad requiere. Estos tres problemas pueden conducir a ciertas distorsiones en las intuiciones o principios que ellos puedan disparar. Por estas razones, para obtener principios morales justificados y que ellos estén en equilibrio reflexivo con nuestras intuiciones y teorías de fondo, el uso de ejemplos reales como tests formales de validez debe completarse con el de experimentos mentales.

---

ellos para argumentar en favor de una concepción concreta de daño que nos permita lidiar con situaciones en las que podemos afectar de manera negativa a las generaciones futuras (Truccone-Borgogno, 2015, 2018b, 2020).

Es inevitable descansar en ambos tipos de ejemplos, al menos para juzgar el valor moral de la historia en nuestras discusiones sobre justicia.

## **Conclusión**

En este trabajo he argumentado que el uso exclusivo de experimentos mentales es insuficiente para lidiar de manera adecuada con algunos problemas importantes de filosofía práctica, tales como la evaluación del rol moral de la historia en nuestras discusiones de justicia actuales. He mostrado, espero, que el uso de ejemplos hipotéticos es problemático dado que, por su mismo diseño, excluye la posibilidad de que la historia pueda tener algún valor en lo que la justicia requiere en el presente. Desde que favorecen un tipo específico de respuesta en la que la historia no puede tener lugar relevante están en conflicto con la idea de imparcialidad a la que deben servir. Para evitar este problema y poder evaluar si, de hecho, la historia debe influir en lo que la justicia requiere en el presente, he argumentado que el uso de ejemplos hipotéticos debe complementarse con la narración de casos reales u ostensivos. Esto es así dado que estos últimos pueden dar lugar a nuevos juicios morales considerados que también están a la espera de ser equilibrados reflexivamente a los fines de poder justificar principios normativos adecuados.<sup>15</sup>

---

15 Debo agradecer muy especialmente a Lucas Misseri y Anna Bugajska por sus comentarios y objeciones, sin las cuales no hubiese podido mejorar este capítulo. Debo agradecer también a Lukas H. Meyer, Tim Waligore y Romina Rekers por diversas discusiones sobre el asunto del valor moral de la historia. Este trabajo es parte de una investigación realizada en el proyecto *Supersession of Historical Injustice and Changed Circumstances*, financiada por el Fondo Austríaco para la Ciencia (FWF) bajo el subsidio para investigación P 30084.

## Referencias bibliográficas

- Bonfiglio, J.; Salvia, A. y Vera, J. (2020). *Deterioro de las Condiciones Económicas de los Hogares y Desigualdades Sociales en tiempos de Pandemia*. Observatorio de la Deuda Social Argentina, Universidad Católica Argentina.
- Brownlee, K. y Stemplowska, Z. (2017). Thought Experiments. En Blau, A., *Methods in Analytical Political Theory* (pp. 21–44). Cambridge University Press.
- Daniels, N. (2016). Reflective Equilibrium. En Zalta Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy*. (Summer 2020 Edition), <https://plato.stanford.edu/archives/sum2020/entries/reflective-equilibrium>
- De la Fuente, D. (1898). *Segundo Censo de la República Argentina*. Taller Tipográfico de la Penitenciaría Nacional.
- Iverson, D. (2000). Political Community and Historical Injustice. *Australasian Journal of Philosophy*, 78(3), 360–373.
- Jamieson, D. (1993). Method and Moral Theory. En Singer P. (Ed.), *A Companion to Ethics* (pp. 476–487). Blackwell Publishing.
- Lyons, D. (1981). The New Indian Claims and Original Rights to Land. En Paul, J. (Ed.), *Reading Nozick: Essays on Anarchy, State and Utopia* (pp. 355–379). Rowman & Littlefield.
- Meyer, L.H. (2001). Transnational Autonomy: Responding to Historical Injustice in the Case of the Roma Peoples. *International Journal on Minority and Group Rights*, 8(2/3), 263–301.
- Meyer, L.H. (2004). Surviving Duties and Symbolic Compensation. En Meyer, L. (Ed.), *Justice in Time: Responding to Historical Injustice* (pp. 173–183). Nomos Verlagsgesellschaft.
- Meyer, L.H. (2006). Reparations and Symbolic Restitution. *Journal of Social Philosophy*, 37(3), 406–422.
- Meyer, L.H. (2020). Intergenerational justice. En Zalta, E. (Ed.), *The Stanford Encyclopaedia of Philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/justice-intergenerational/>
- Nagy, M. (2012). *Tradiciones situadas, usos del pasado y devenir indígena*. Vols. 1 y 2. Tesis Doctoral, Universidad de Buenos Aires. <http://repositorio.filo.uba.ar/handle/filodigital/2822>
- Nagy, M. (2014). Después del fin. Sometimiento, proletarianización y rearticulación comunitaria indígena en Buenos Aires. En Lanata, J.L. (Comp.), *Prácticas Genocidas y Violencia Estatal en Perspectiva Transdisciplinaria* (pp. 96–133). IIDyPCa–CONICET–UNRN.
- Nino, C. (1991). *The Ethics of Human Rights*. Oxford Clarendon Press.
- Parfit, D. (1984). *Reasons and Persons*. Clarendon Press.
- O’Neill, O. (1986). The Power of Example. *Philosophy*, 61, 5–29.
- Rawls, J. (1971). *A Theory of Justice, original edition*. Harvard University Press.
- Shue, H. (2009). Making exceptions. *Journal of Applied Philosophy*, 26(3), 307–322.
- Truccone-Borgogno, S. (2015). Derecho moral y el problema de la no-identidad: apuntes sobre el principio de daño. *Doxa*, (38), 473-499.
- Truccone-Borgogno, S. (2018a). Estudio introductorio al equilibrio reflexivo y a la función de los ejemplos en la investigación moral. En *XVII Anuario del Centro de Investigaciones Jurídicas y Sociales de la Facultad de Derecho de la UNC* (pp. 687–709). La Ley.
- Truccone-Borgogno, S. (2018b). Consideraciones sobre la fuerza de las razones en contra de dañar. *Crítica: Revista Hispanoamericana de Filosofía*, 50(149), 31–57.

- Truccone-Borgogno, S. (2020). El problema de la no-identidad: cuatro soluciones posibles. *Ideas y Valores: Revista Colombiana de Filosofía*, (172), 57-80.
- Truccone-Borgogno, S. (2021a) *The Temporal Dimension of Justice*. Doctoral Dissertation, University of Graz.
- Truccone-Borgogno, S. (2021b) La Conquista del Desierto, Confianza y el principio de Proximidad. *Análisis Filosófico*, 4(1), 7–36.
- Truccone-Borgogno, S. (2022a). The Supersession Thesis, Climate Change, and the Rights of Future People. en *Superseding historical injustice? New critical assessments*, Special Issue edited by Lukas H. Meyer and Timothy Waligore. *Critical Review of International Social and Political Philosophy (CRISPP)*, 25(3), 364–379.
- Truccone-Borgogno, S. (2022b) Responding to historical injustices: collective inheritance and the moral irrelevance of group identity. *European Journal of Political Theory* (pp. 1-20), Early view. <https://doi.org/10.1177/14748851221100094>
- Truccone-Borgogno, S. (2022c). Climate Justice and the Duty of Restitution. *Moral Philosophy and Politics* (1-22). Ahead of Print. <https://doi.org/10.1515/mopp-2021-0071>
- Vernon, R. (2003). Against Restitution. *Political Studies*, 51, 342–557.
- Wenar, L. (2006). Reparations for the Future. *Journal of Social Philosophy*, 37(3), 396–405.
- Waldron, J. (1992a). Historic Injustice: Its Remembrance and Supersession. En Oddie, G. y Perrett, R. (Ed.), *Justice, Ethics, and New Zealand Society* (pp. 139–170). Oxford University Press.
- Waldron, J. (1992b). Superseding Historic Injustice. *Ethics*, 103(1), 4–28.
- Waligore T. (2016). Rawls, Self-Respect, and Assurance. How Past Injustice Changes What Counts as Justice. *Politics, Philosophy & Economics*, 15, 42–66.
- Waligore T. (2018). Redress for Colonial Injustice: Structural Injustice and the Relevance of History. *Global Justice. Theory Practice Rhetoric*, 11, 15–28.

## 12. Desacuerdos morales

David Martínez Zorrilla

El análisis conceptual es uno de los rasgos más característicos y uno de los pilares centrales de lo que puede denominarse la *tradición analítica* en la Filosofía. Todo análisis y clarificación conceptual es, al menos en parte, una actividad de carácter estipulativo, lo que suele ser un motivo de crítica por parte de otras concepciones o enfoques de la actividad filosófica, en la medida en que la estipulación permite tornar algo en «correcto» o «verdadero» por definición del propio sujeto que la propone. Pero aunque este sea el caso, toda estipulación también puede (y probablemente *debe*) ser evaluada mediante criterios sustantivos externos a la propia estipulación, entre los que ocupa un lugar destacado el de la compatibilidad o coherencia con las intuiciones preanalíticas de los sujetos que participan en el discurso. Se trata de encontrar, usando la famosa expresión de J. Rawls, un «equilibrio reflexivo» entre tales intuiciones y el análisis propuesto, de modo que los conceptos sean reconocidos como tales por los participantes, al tiempo que se gana rigor, precisión y claridad en el uso de los mismos.

Ningún análisis conceptual parte del vacío y siempre precisa de algunas asunciones previas como punto de partida. En el presente caso, serían fundamentalmente de las siguientes:

- 1) Para simplificar el análisis, en este trabajo se usarán los términos «moral» y «ética» como sinónimos,<sup>1</sup> esto es, indistintamente, aunque debe tenerse en cuenta que algunos autores trazan una distinción, utilizando uno de los términos para referirse a los juicios sobre aquello que es correcto/incorrecto y/o debido/permitido/prohibido hacer (o a las creencias y/o actitudes que un determinado grupo mantiene acerca de lo que es (in)correcto o (in)debido, lo que es algo distinto), y el otro para referirse al análisis filosófico sobre lo anterior (es decir, como una rama de la filosofía práctica). No será así en el presente caso.
- 2) Se parte también de una división o clasificación ampliamente asumida según la cual es posible diferenciar al menos tres distintos niveles o dimensiones en el análisis de la moralidad: a) la *moral social* o *ética descriptiva*; b) la *ética normativa*; y c) la *metaética*. Dicho de manera extremadamente sin-

---

1 A fin de cuentas, y aunque no sea demasiado relevante para el análisis, desde un punto de vista etimológico, tanto *mores* en latín como *ethos* en griego pueden traducirse por «costumbres» y hacen referencia fundamentalmente a lo mismo.

tética, la moral social hace referencia a las creencias y actitudes que, de hecho, un determinado grupo social o colectivo de personas mantiene acerca de lo que es moralmente (in)correcto o (in)debido. La ética normativa se refiere propiamente a los juicios éticos, es decir, a lo que se considera como un comportamiento moralmente (in)correcto o (in)debido, si bien normalmente también suelen incluirse en este ámbito las teorías normativas, esto es, aquel conjunto de principios o criterios que son los que determinan que un comportamiento sea considerado precisamente como éticamente (in)debido o (in)correcto.<sup>2</sup> Finalmente, la metaética incluye la reflexión filosófica acerca del significado de los conceptos y el lenguaje moral (por ejemplo, qué quieren decir de manera precisa expresiones como «debo moralmente hacer X», o «Y es inmoral»), lo que impacta y se conecta con otros múltiples aspectos como la fundamentación o justificación de los juicios morales, o la epistemología moral. Como veremos, la distinción entre estos niveles tiene un importante impacto de la determinación precisa de qué significa que existe un desacuerdo en materia moral.

- 3) En tercer lugar, para poder analizar en qué consiste un desacuerdo *moral* resulta preciso partir de alguna noción más general de lo que es un *desacuerdo*. Por lo pronto, la idea de desacuerdo parece estar unida a la existencia de algún tipo de incompatibilidad entre lo que podrían denominarse de manera genérica como «posiciones» que distintos agentes o sujetos mantienen acerca de algo. Tales «posiciones» pueden consistir tanto en proposiciones (significados de enunciados descriptivos) como en normas (si se trata de enunciados deónticos). Así, podrían considerarse como ejemplos de desacuerdo tanto la afirmación del sujeto A «El gato está sobre el sofá» respecto de la del sujeto B «El gato no está sobre el sofá», como el enunciado del sujeto C «El comportamiento x está prohibido» respecto del sujeto D «El comportamiento x está permitido». Resulta preferible el uso del término más neutral «incompatibilidad» en lugar del más preciso «inconsistencia» (incompatibilidad lógica o contradicción) para evitar entrar en el debate acerca de la (in)existencia de relaciones lógicas entre enunciados normativos o deónticos.<sup>3</sup>

---

2 Así, por ejemplo, teorías como el utilitarismo moral o la deontología kantiana suelen ser consideradas como concepciones de ética normativa, aunque en sentido estricto más que establecer qué debemos hacer lo que hacen es establecer los criterios o condiciones que *determinan* o justifican que algo sea moralmente debido o correcto. En este sentido, quizá sería más preciso y coherente considerar tales teorías como pertenecientes al ámbito de la metaética, si se considera que esta incluye la reflexión acerca de la justificación o fundamentación de los juicios morales. En cualquier caso, no entraremos en este debate.

3 Cualquier propuesta de análisis sobre las relaciones lógicas entre normas debe enfrentarse al llamado «dilema de Jørgensen», por ser este el autor que lo expuso en un famoso artículo (ver Jør-

Nótese también que en la propuesta de noción genérica de desacuerdo se hace referencia a la intervención de varios sujetos o agentes. Ello se debe a que incluso en un nivel intuitivo resulta bastante extraño afirmar que un mismo sujeto está en «desacuerdo» consigo mismo en la hipótesis de que sostenga posiciones incompatibles. En estos casos suele afirmarse más bien que el sujeto es inconsistente (o que sostiene posiciones o afirmaciones inconsistentes/incompatibles), o que se contradice a sí mismo. Por ello, la presencia de al menos dos sujetos parece ser un elemento o requisito necesario del propio concepto de «desacuerdo».<sup>4</sup>

En este punto, podría decirse en una primera aproximación que un «desacuerdo moral» consiste en cualquier desacuerdo (esto es, una situación en la que diversos sujetos mantienen sendas posiciones incompatibles) acerca de algún aspecto o extremo que pueda considerarse como de tipo o con contenido moral. Ahora bien, en función de cómo se entienda que algún asunto o cuestión es «de tipo moral», el desacuerdo moral puede significar cosas distintas, o dicho de otro modo, podemos hablar de distintos sentidos o conceptos de «desacuerdo moral». Para poder diferenciarlos y tener una idea más precisa de los mismos, resultará útil detenerse en cada una de las distintas dimensiones de «la moral» antes indicadas: la ética descriptiva o moral social, la ética normativa y la metaética.

- 4) En relación con el concepto de *error*, este parece requerir conceptualmente de la existencia de pautas, criterios o reglas de corrección, en la medida en que aquel se concibe como lo que se aparta o no coincide con una «respuesta correcta». Tales pautas, criterios o reglas pueden ser de muy diverso tipo, en función del ámbito (reglas conceptuales como las del álgebra o la lógica deductiva, principios normativos, etc.),<sup>5</sup> si bien

---

gensen, 1938:288–298). En síntesis, dado que tradicionalmente el concepto de *consecuencia lógica* se ha definido en términos de verdad y suele afirmarse que las normas no son verdaderas ni falsas, entonces o bien no se plantean relaciones lógicas entre normas, o bien la lógica no puede basarse únicamente en la verdad.

4 Es remarcable el contraste en este punto respecto de los conflictos morales. Habitualmente, en el debate filosófico, estos suelen conceptualizarse desde la perspectiva de un único agente o sujeto, que no puede satisfacer todas las exigencias morales que, al menos *prima facie*, le son aplicables, en la medida en que la satisfacción de alguna(s) de ella(s) impide la satisfacción del resto. Para una visión general del debate acerca de los conflictos y dilemas morales. Ver, por ejemplo, Gowans (1987) y Mason (1996).

5 Estos criterios no necesariamente implican que siempre exista una *única* respuesta correcta en cada caso, por lo que la presencia de un desacuerdo no supone que al menos uno de los sujetos sostenga una posición errónea. Un ejemplo muy simple sería el siguiente: si al preguntar por el resultado de la raíz cuadrada de «4», un sujeto A responde «+2» y un sujeto B «-2», a pesar de las distintas respuestas, ambos estarán en lo correcto, pues de acuerdo con las reglas del álgebra, tanto el cuadrado de «+2» como el de «-2» dan como resultado 4.

cuando se trata de proposiciones asertivas hay un criterio que destaca especialmente: el de la *verdad*.<sup>6</sup> Con todo, es importante diferenciar entre el aspecto conceptual de la verdad (que la proposición *p* sea verdadera o falsa) del aspecto epistemológico (cómo saber o determinar si la proposición *p* es verdadera o falsa). Siempre que se acepten los principios de bivalencia y tercero excluido, toda proposición *p* o bien es verdadera o bien es falsa, pero puede que no resulte posible determinar su valor de verdad. Esta imposibilidad puede ser estrictamente técnica (por no disponer de los medios o mecanismos que permitirían verificar o contrastar la proposición),<sup>7</sup> pero también puede ser lógica o conceptual.<sup>8</sup> Aplicado al discurso moral, el «error moral» consistiría en cualquier desviación o incompatibilidad con respecto a los criterios o estándares de «corrección moral» que puedan establecerse.

### Los desacuerdos morales en la ética descriptiva

Como se ha indicado anteriormente, la ética descriptiva o moral social puede concebirse como el conjunto de afirmaciones acerca de las creencias y actitudes mayoritariamente asumidas de un grupo social acerca de cuestiones morales o con relevancia moral. Ejemplos de afirmaciones de este tipo serían, entre muchas otras, «la esclavitud no era considerada una institución injusta en la Antigüedad», o «la mayoría de los miembros de la comunidad *c* considera inmoral el aborto». Tales afirmaciones expresan proposiciones descriptivas, que como tales pueden resultar verdaderas o falsas en función de cómo sean las creencias y actitudes que de hecho tienen los miembros del grupo de referencia. Un aspecto destacable de las proposiciones de la moral social es que estas son necesariamente *relativas*, en la medida en que su verdad o falsedad dependerá del grupo social en cuestión (con existencia temporal y espacial determinada) respecto del que se formula la afirmación. De ese modo, una afirmación como «la esclavitud no es injusta» puede expresar una proposición

---

6 No entraré aquí en el debate sobre las diversas teorías filosóficas acerca de la verdad, como la correspondencia, el desentrecorrido o la coherencia.

7 Por ejemplo, a afirmación «hay una taza de café más allá de la órbita de Plutón» es necesariamente verdadera o falsa, aunque resulte técnicamente imposible determinar su valor de verdad porque no existen telescopios lo suficientemente potentes para localizar un objeto tan pequeño y lejano, ni los instrumentos más sensibles disponibles permiten detectar las alteraciones gravitacionales generadas por un objeto con tan poca masa.

8 Una proposición como «Dios existe» es necesariamente verdadera o falsa, pero si se asume la tesis kantiana de que las fronteras del conocimiento están configuradas por los límites de la experiencia posible, resulta conceptualmente imposible determinar la verdad o falsedad de la proposición, al referirse a una cuestión metafísica.



verdadera si se refiere a la sociedad romana del siglo I, o una proposición falsa si se vincula a una sociedad occidental del siglo XXI, mientras que «la pena de muerte es inmoral» probablemente expresa una proposición verdadera si se refiere a la sociedad sueca actual, mientras que posiblemente sea falsa si se refiere a la sociedad norteamericana. Ello también implica que dos proposiciones que en abstracto resultan contradictorias o inconsistentes, como «la eutanasia es inmoral» y «la eutanasia no es inmoral» no son necesariamente contradictorias y pueden ser ambas verdaderas, si se refieren a grupos sociales distintos.<sup>9</sup>

Para poder hablar de un desacuerdo en el ámbito de la moral social, se requiere entonces que *dos o más agentes afirmen sendas proposiciones acerca de las creencias y actitudes morales de un mismo grupo social de referencia que no solo resulten distintas, sino mutuamente excluyentes* (es decir, que una de ellas implique lógicamente la negación de la otra).

Dado que los contenidos de la ética descriptiva están formados por proposiciones asertivas sobre ciertas realidades empíricas que actúan como referencia de las mismas, el criterio primordial para poder hablar de acierto o error es el de la verdad, y los mecanismos para determinar epistémicamente la verdad o falsedad de las proposiciones serán los propios de disciplinas como la antropología social o la sociología (si se trata de sociedades o comunidades actuales), o los de la investigación histórica, si se trata de sociedades pasadas.

Según los principios de bivalencia y del tercero excluido, si dos proposiciones son contradictorias, necesariamente la verdad de una de ellas implicará la falsedad de la otra, y viceversa. Pero esto no implica necesariamente que uno de los sujetos «tenga la razón», al menos en un sentido relevante de la expresión. Imaginemos el ejemplo siguiente: Supongamos que un sujeto A afirma que «aproximadamente un 10 % de los miembros de la sociedad s aprueba los matrimonios poligámicos, y el restante 90 % está en contra», mientras que

---

<sup>9</sup> Incluso, es posible que las creencias morales mayoritariamente asumidas en una comunidad determinada sean inconsistentes, y en tal caso la afirmación de la existencia de esa contradicción no sería en sí inconsistente, sino una proposición verdadera, puesto que una proposición  $p$  que afirma que un determinado conjunto de proposiciones  $\Sigma$  es inconsistente no es en sí una afirmación inconsistente, al ser metalingüística respecto de tal conjunto. Un ejemplo que puede ilustrar esta situación es el de los resultados de la investigación realizada por el MIT conocida como *Moral Machine* ([www.moralmachine.net](http://www.moralmachine.net)), con el objetivo de descubrir las intuiciones morales que de hecho tienen las personas, para su posible implementación en los algoritmos de conducción autónoma de vehículos, cuando se enfrentan a una situación en la que resulta imposible evitar un accidente y causar víctimas. Los experimentos muestran que mientras que habría una mayoría de personas que se inclinaría, siguiendo parámetros utilitaristas, por la opción que implique el menor número de víctimas, al mismo tiempo una clara mayoría no estaría dispuesta a adquirir o utilizar un vehículo programado para «sacrificar» a sus ocupantes, aunque esa fuera la alternativa menos dañina en términos utilitaristas.

un sujeto B sostiene que «aproximadamente un 15 % de los miembros de la sociedad S aprueba los matrimonios poligámicos, y el restante 85 % está en contra». Las afirmaciones de A y B se excluyen mutuamente, puesto que si es cierto que el 10 % está a favor y el 90 % en contra, necesariamente es falso que el 15 % esté a favor y el 85 % en contra (y si es falso que el 10 % está a favor, es necesariamente verdadera la proposición «no es el caso que el 10 % de los miembros de S están a favor de la uniones poligámicas», que es una consecuencia lógica de la afirmación de B). El mismo razonamiento puede hacerse con relación a B respecto de A. Supongamos que, por hipótesis, en realidad un 20 % de los miembros de S está a favor del reconocimiento de las uniones poligámicas y el 80 % está en contra. En tal caso, parecería perfectamente legítimo decir que tanto A como B están equivocados o en un error.<sup>10</sup>

### **Los desacuerdos morales en el ámbito de la ética normativa**

En una aproximación muy general, puede decirse que la ética normativa está formada por el conjunto de los *juicios éticos* o *morales*, entendiéndolo por éstos los enunciados que califican un determinado comportamiento o estado de cosas como moralmente correcto o incorrecto, bueno o malo, justo o injusto, debido o indebido, etc. Por tanto, se trata del ámbito del discurso humano relativo cómo hay que actuar y comportarse para que dicho comportamiento pueda considerarse como ajustado a los parámetros morales (esto es, como éticamente correcto, adecuado o al menos permitido), como en los enunciados «es un deber moral cumplir con los compromisos adquiridos», o «es inmoral causar voluntariamente un daño a un inocente». Se trata probablemente del sentido más habitual que los individuos normalmente asocian a los términos «ética» o «moral».

El adjetivo «normativa» remarca el hecho de que al menos en principio estos juicios tienen una función de guiar la conducta o determinar pautas de comportamiento que se consideran como éticamente correctas (pues parece existir algún tipo de contradicción pragmática o *performativa*, por usar esta expresión probablemente no muy adecuada en español, entre afirmar «es moralmente correcto hacer X» y en cambio no llevarlo a cabo), aunque exis-

---

<sup>10</sup> El ejemplo siguiente resultaría aún más evidente: si A afirma «la nieve es roja» y B afirma «la nieve es verde», la afirmación de A implica la negación de la de B y viceversa, pero ambas proposiciones son falsas en la medida en que la nieve no es ni roja ni verde, sino blanca. Por tanto, la existencia de un desacuerdo en una cuestión empírica no supone que uno de los implicados esté en lo cierto.

te debate acerca de si los juicios morales son descriptivos o prescriptivos. En cualquier caso, este es un aspecto cuyo análisis pertenece al ámbito de la metaética, de modo que no será abordado por el momento.

Debe tenerse en cuenta también que, debido a problemas de ambigüedad, es posible que un mismo enunciado pueda ser interpretado, según las circunstancias y el contexto, como un juicio moral o una proposición de ética descriptiva. Un enunciado como «el aborto es inmoral» puede entenderse, al menos en ciertos contextos, tanto como un juicio moral genuino que califica esta práctica como moralmente incorrecta y de algún modo pretende que no se lleve a cabo, o como una mera descripción acerca de las creencias y actitudes morales de un grupo social (ética descriptiva). Es importante destacar que no podemos hablar de la existencia de un desacuerdo moral entre distintos sujetos si estos se mueven en distintos planos del discurso. Si un sujeto A sostiene «el aborto es inmoral», mientras que otro sujeto B afirma «el aborto es éticamente correcto en determinados casos», y uno de los enunciados se interpreta como una proposición de ética descriptiva y el otro como un juicio moral, no habría realmente ningún desacuerdo entre ambos. Es más, un mismo sujeto podría sostener ambas afirmaciones sin contradecirse.

Partiendo de esta base, de modo genérico un «desacuerdo moral» en el ámbito de la ética normativa consistiría en toda situación en la que dos o más sujetos sostienen o defienden sendos juicios morales incompatibles. No obstante, una complicación que se plantea es que, como se señaló más arriba, es bastante habitual incluir también en el ámbito de la ética normativa a aquellas *teorías normativas* (como por ejemplo el utilitarismo o el deontologismo kantiano) que sirven para determinar o justificar qué comportamientos o estados de cosas son moralmente correctos. En este punto podría argumentarse que, estrictamente hablando, dichas teorías pertenecerían al ámbito de la metaética, porque se refieren a las condiciones para la *justificación* de los juicios morales (por ejemplo, según el utilitarismo clásico, un comportamiento sería éticamente correcto en la medida en que maximice la felicidad del mayor número, mientras que para un kantiano la justificación dependería de la satisfacción de las exigencias impuestas por el imperativo categórico), y no a los juicios morales mismos.

Si se considera que tales teorías normativas forman parte del ámbito de la ética normativa, debería en consecuencia afirmarse la existencia de un desacuerdo moral entre dos o más sujetos por el mero hecho de sostener teorías normativas incompatibles, independientemente de que los juicios morales que sean justificados o respaldados por dichas teorías sean o no compatibles. En cambio, si el ámbito de la ética normativa se circunscribe estrictamente a

los juicios morales, en muchos casos no se plantearán desacuerdos a pesar de que los sujetos mantengan diferentes teorías normativas.<sup>11</sup>

Los juicios éticos pueden ser generales o particulares (aunque se trataría de una distinción gradual y por tanto sujeta a cierta dosis de vaguedad). *Grosso modo*, los juicios generales se corresponderían con los «grandes» principios morales, o los considerados más básicos, mientras que los juicios particulares se referirían a acciones o comportamientos concretos bajo las determinadas circunstancias empíricas de un caso individual, y que serían, al menos teóricamente, concreciones o aplicaciones de los principios o juicios generales al caso concreto. Así, para un utilitarista clásico un juicio moral general podría ser «es moralmente obligatorio (correcto, debido, etc.) actuar de modo que se procure la mayor felicidad para el mayor número», mientras que para un kantiano uno de estos juicios podría ser «es inmoral (moralmente prohibido, incorrecto, etc.) infligir voluntariamente un daño a una persona inocente». Debe destacarse que no existe conflicto ni incompatibilidad lógica alguna entre ambos juicios generales. Si se interpretan como normas categóricas, la manera más sencilla de representarlas formalmente sería como «O(p)» y «Ph(q)», lo que no constituye ninguna contradicción o antinomia en absoluto. Cuando se aplican a casos individuales para obtener juicios particulares, en muchas ocasiones tampoco se planteará incompatibilidad alguna (en muchos casos, actuar de modo que se maximice la felicidad del mayor número no implicará infligir un daño a personas inocentes). La incompatibilidad (y por tanto el desacuerdo) surgirá cuando debido a las circunstancias del caso a decidir los juicios éticos particulares resultantes sean mutuamente excluyentes: si en las circunstancias *C* el comportamiento *c* es aquel que promueve la mayor felicidad al mayor número al tiempo que supone infligir un daño a un inocente, el sujeto partidario de la concepción utilitarista defenderá el juicio moral (particular) «es moralmente correcto (obligatorio, bueno, debido, etc.) realizar *c*», mientras que el sujeto partidario de una concepción kantiana de la ética defenderá el juicio moral particular «es moralmente incorrecto (prohibido, malo, etc.) realizar *c*», planteándose en consecuencia un desacuerdo moral.

A fin de poder proporcionar un concepto más preciso de «desacuerdo moral» en el ámbito de la ética normativa, que vaya más allá de la afirmación ge-

---

11 Incluso, también puede ocurrir que surjan desacuerdos éticos por incompatibilidad entre los juicios morales aún entre sujetos que compartan una misma teoría normativa. Por ejemplo, dos sujetos que compartan una concepción utilitarista clásica pueden disentir acerca de cuál de entre dos cursos de acción posibles pero excluyentes maximiza el bienestar del mayor número (lo que es un desacuerdo empírico pero que desemboca en un desacuerdo de juicios morales), o dos sujetos que sostengan una concepción kantiana pueden tener opiniones diversas acerca de si un determinado comportamiento supone tratar a una persona solamente como un medio y no como un fin en sí mismo.

nérica de que dos o más sujetos sostienen juicios éticos incompatibles, resulta de gran interés recurrir a uno de los análisis más rigurosos y completos del concepto de «deber moral» (u «obligación moral»), como es el desarrollado por Michael J. Zimmerman.<sup>12</sup> En su análisis, el autor recurre a algunas ideas clave que le sirven de base para su propuesta, como las de *comportamiento personalmente opcional*, y la *de mundo accesible*. Para el autor, una obligación o deber moral puede definirse (aunque todavía no formalmente) como aquel comportamiento *personalmente opcional* que conduzca al *mundo accesible* con un valor deóntico superior al del resto de alternativas de comportamiento personalmente opcionales.

- a) Que un comportamiento sea *personalmente opcional* significa que el agente puede (empíricamente) tanto realizarlo como abstenerse de realizarlo. Por tanto, excluye todas aquellas acciones y omisiones que resulten imposibles o necesarias, y con ello asume el usualmente denominado «principio kantiano» o «principio voluntarista» del «debe» implica «puede» («ought» implies «can»). Por ello nos encontramos ante conductas que serían una «opción personal» para el agente, en la medida en que el realizarlas o no responderá a una decisión libre (en sentido empírico, no deóntico) de este.
- b) Por otra parte, un *mundo accesible* es un estado de cosas al que el agente puede llegar a través de la realización de una conducta (acción u omisión), y que se caracteriza parcialmente por el propio hecho de ser un mundo en el que el agente realiza tal comportamiento. La conducta del agente determinará entonces que se acceda a dicho mundo, que es un mundo en el que, junto a muchas otras propiedades, se caracteriza por el hecho de que el agente se ha comportado de cierto modo.

Lo que sea considerado como moralmente obligatorio dependerá de la evaluación y comparación del valor deóntico (conforme a una teoría determinada, como el utilitarismo o la ética kantiana) de los distintos mundos accesibles abiertos al agente como resultado de comportamientos personalmente opcionales. A partir de lo anterior, Zimmerman propone una definición más rigurosa del concepto de «obligación moral», según la cual «el agente *S* debe, en el momento *T*, hacer *A* en *T*, si y sólo si: a) *S* puede (empíricamente) hacer *A* en *T*; b) *S* puede (empíricamente) abstenerse de hacer *A* en *T*; y c) todo mundo accesible en el que *S* no hace *A* en *T* es deónticamente inferior a algún mundo en el que *S* hace *A* en *T*»<sup>13</sup> (1996:26). En esta definición, un juicio moral se en-

---

12 Ver Zimmerman (1996).

13 A partir de la definición anterior es posible derivar un concepto de «permisión moral», según el cual: «el agente *S* puede (deónticamente), en el momento *T*, hacer *A* en *T*, si y sólo si: a) *S* puede (empíricamente) hacer *A* en *T*; b) *S* puede (empíricamente) abstenerse de hacer *A* en *T*; y c)

tendería como un conjunto o conglomerado de afirmaciones que hacen referencia tanto a la posibilidad empírica de un agente para realizar un determinado comportamiento, como al valor que en términos éticos se atribuye a tal comportamiento.

El hecho de comparar alternativas de acción en términos del *valor deóntico* que tienen los mundos o estados de cosas a que aquellas dan lugar puede llevar a pensar que esta reconstrucción solamente funciona con respecto a las teorías consecuencialistas, o más específicamente utilitaristas, en las que el valor moral de las conductas depende directamente del grado de consecución de algo que se considera intrínsecamente valioso. Pero como el propio autor señala (en mi opinión acertadamente), su propuesta de análisis conceptual también puede aplicarse a las teorías de tipo deontológico, en tanto que el «valor moral» de los mundos o estados de cosas en los que se satisfacen los principios morales es superior al de aquellos otros en el que no son satisfechos. Por tanto, la propuesta de Zimmerman puede considerarse neutral o «agnóstica» con respecto a las teorías normativas sustantivas.<sup>14</sup>

Partiendo de este análisis del concepto de «deber moral», puede afirmarse que existe un desacuerdo moral en sede de la ética normativa cuando diferentes sujetos o agentes mantienen sendas posiciones distintas e incompatibles (mutuamente excluyentes o contradictorias) acerca de uno o varios de los elementos que configuran la afirmación de la existencia de un deber moral, esto es, o bien acerca de la existencia o no de un determinado mundo personalmente accesible en el que se hace o no hace A en T', y/o acerca del valor deóntico asignado a un determinado mundo accesible en el que se hace o no hace A

---

*ningún mundo accesible en el que S no hace A en T' es deónticamente superior a algún mundo en el que S hace A en T'» (Zimmerman, 1996:32).*

14 Otra cuestión que puede plantear dudas acerca de la adecuación de la propuesta es que, en la medida en que se define el «deber moral» en términos de la existencia de comportamientos personalmente opcionales y mundos accesibles con determinado valor deóntico, eso parece implicar que los juicios morales son descriptivos, lo que podría plantear problemas con respecto a aquellas concepciones metaéticas conforme a las cuales los juicios éticos son prescriptivos o normativos, y no descripciones, ya sean de estados de cosas o de propiedades, sean estos naturales o no naturales. En mi opinión, una vía para intentar superar esta dificultad consistiría en acudir a la distinción introducida por G.H. Von Wright entre normas y proposiciones normativas (ver Von Wright, 1963:93), conforme a la cual es posible diferenciar entre un enunciado deóntico o normativo y el correspondiente enunciado descriptivo o proposicional que afirma la existencia de dicha norma. La propuesta de Zimmerman podría entenderse como una reconstrucción de la proposición normativa que afirma (describe) la existencia de un juicio moral normativo correspondiente que establece lo moralmente obligatorio/permitido/prohibido. En este sentido, la reconstrucción del autor haría referencia a qué significa que *existe* una obligación o deber moral, o, si se prefiere, a qué condiciones deben estar presentes para que sea verdadera la proposición que afirma que existe la obligación moral para S de hacer A en T'.

en  $\tau'$ . Por tanto, explicado en términos menos formales, el desacuerdo puede responder tanto a un disenso acerca de la posibilidad empírica para un agente de realizar una determinada conducta,<sup>15</sup> como a una distinta valoración en términos morales de tal conducta.

¿Qué supone «estar en un error» en el ámbito de la ética normativa? Al ser los juicios morales expresiones acerca de cómo hay que comportarse, para poder determinar si lo que establecen dichos juicios es correcto o se trata de un error, es necesario analizar la *justificación* o *fundamentación* de tales juicios, lo cual trasciende el propio ámbito de la ética normativa y corresponde al campo de la metaética. Los juicios morales serán «correctos» o «erróneos» en función de los criterios establecidos por la teoría o concepción metaética que se maneje, por lo que se retomará esta cuestión más adelante, en el próximo apartado. Con todo, existe un aspecto en el que sí que puede ofrecerse ya una respuesta: si se entiende, siguiendo a Zimmerman, que las proposiciones empíricas acerca del carácter personalmente opcional para el agente *s* de un comportamiento forman parte de los juicios de deber moral, el criterio determinante para poder hablar de acierto o error en este punto será el de la *verdad* de las proposiciones referentes a la (im)posibilidad de realizar tal comportamiento para el agente *s*.

## Los desacuerdos morales en el ámbito de la metaética

De forma muy general, la metaética puede concebirse como la reflexión filosófica acerca de los distintos aspectos que conforman el discurso moral.<sup>16</sup> En lugar de centrarse directamente en responder a la pregunta «¿qué debo (moralmente) hacer?», analiza cuestiones como el significado y estructura de los juicios morales (¿qué significa un enunciado como «*x* es moralmente co-

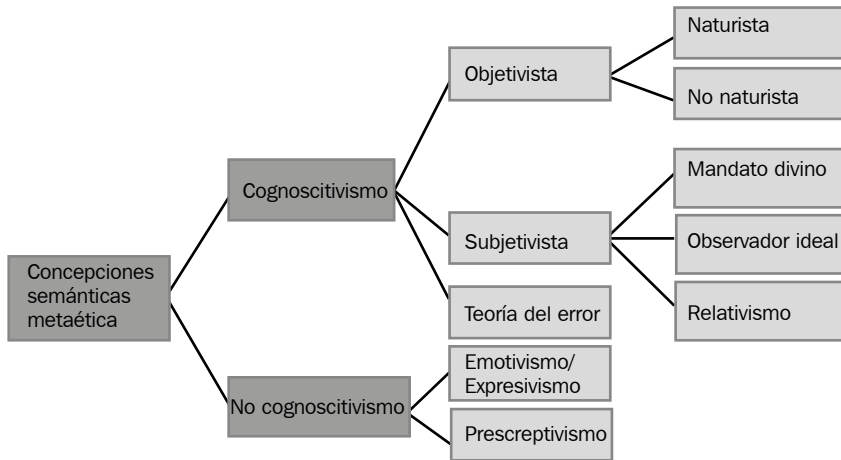
---

<sup>15</sup>Esta propuesta solo funciona en la medida en que se asuma como un presupuesto para la existencia de un deber u obligación moral el principio voluntarista del «debe implica puede», extremo que no es pacífico en la literatura. El análisis de dicho principio queda fuera de los límites de este trabajo, si bien parece razonable considerar que si la moral tiene como objetivo la guía de la conducta humana, entonces debe referirse a comportamientos que sea posible realizar u omitir, ya que no parece tener mucho sentido ordenar o prohibir aquello que es imposible o inevitable hacer. Por tal razón la constatación de que un comportamiento resulta o no optativo para el agente tiene un impacto directo en su posibilidad de valoración como deber moral, y un desacuerdo acerca de esta cuestión (que es estrictamente empírica) puede constituir una fuente de desacuerdo acerca de lo que es moralmente correcto, bueno o debido, al igual que un disenso respecto del valor deóntico asignado a una acción o comportamiento determinados.

<sup>16</sup>Existen bastantes obras de referencia para tener una visión general del debate sobre la metaética. A modo de ejemplo, pueden enumerarse las siguientes: Brandt (1959); Hudson (1970); Darwall, Gibbard y Railton (1997); Van Roojen (2015).

recto»?), su justificación o fundamentación (¿qué es lo que determina aquello que es moralmente bueno, correcto o justo y permite diferenciarlo de lo moralmente malo, incorrecto o injusto?), la epistemología moral (¿es posible conocer, y de qué modo, lo que es moralmente correcto?), la ontología moral (¿existen valores o principios morales objetivos y universales?), etc. En realidad, todas estas cuestiones están estrechamente relacionadas, y la respuesta que se ofrezca a alguna de ellas condiciona en gran medida la de las demás.<sup>17</sup>

En lo que podría considerarse como los inicios de la metaética contemporánea (a principios del siglo xx), es decir, como disciplina filosófica diferenciada y explícita y conscientemente centrada en el análisis del discurso moral, el interés primordial se centraba en el análisis semántico de los juicios morales, y podría decirse que hoy en día todavía es este ámbito el que suele tomarse principalmente en cuenta a la hora de diferenciar y clasificar las distintas corrientes o posiciones en el campo de la metaética. En este sentido, y aunque estrictamente hablando podría decirse que existen tantas concepciones distintas como autores, existe cierto consenso en agrupar las diferentes teorías o corrientes metaéticas en las siguientes categorías principales:



<sup>17</sup> Por ejemplo, para un autor que considere que semánticamente los juicios morales son enunciados descriptivos acerca de propiedades o estados de cosas naturales (observables empíricamente), la manera de discernir entre los juicios morales correctos e incorrectos será mediante la verificación empírica, a través de la observación, de modo que solo estarán justificados aquellos que se ajusten a proposiciones verdaderas acerca de ciertos hechos naturales.



Sin entrar en detalles, puede verse que la distinción fundamental se sitúa entre aquellas concepciones denominadas «cognoscitivistas» y aquellas otras «no-cognoscitivistas». Las primeras, pese a sus grandes diferencias, tienen en común el hecho de concebir los juicios morales como enunciados descriptivos, y que por tanto expresan proposiciones con valores de verdad. Las diferencias entre las distintas variedades se encuentran precisamente en qué es aquello que se supone que describen. Por su parte, las concepciones metaéticas no-cognoscitivistas sostienen en cambio que los juicios morales no consisten (o no solo, o no principalmente) en proposiciones descriptivas, sino que se trata de enunciados cuya función o fuerza ilocucionaria no consiste en describir o transmitir información de algún tipo.

Dentro de las posiciones cognoscitivistas, la principal división se establece entre quienes sostienen que aquello que describen los juicios éticos son las creencias y/o actitudes de algún sujeto o conjunto de sujetos (cognoscitivismo subjetivista), o bien algo distinto de lo anterior (cognoscitivismo objetivista). También se pueden considerar como una versión del cognoscitivismo la llamada teoría del error, que tiene ciertas características peculiares que impiden ubicarla fácilmente tanto en las concepciones objetivistas como en las subjetivistas.

Dentro del cognoscitivismo objetivista, pueden distinguirse las posiciones «naturalistas» de las «nonaturalistas». Para las primeras, lo que describen los juicios morales son propiedades o estados de cosas de carácter empírico, observable o natural, a partir de las cuales superviene su calificación ética. Por ello, en último término los juicios sobre (in)corrección moral son «traducibles» a proposiciones sobre determinadas propiedades o circunstancias naturales. Uno de los ejemplos más claros es el del utilitarismo moral (al menos ciertas versiones). Si se considera por ejemplo que el placer o el bienestar del mayor número es lo que otorga valor moral a los comportamientos, se está sosteniendo que el valor moral de una acción superviene a una característica o propiedad natural y empíricamente verificable, y por tanto que una expresión como « $x$  es moralmente bueno» significa básicamente « $x$  produce placer al mayor número», que es una proposición descriptiva sobre hechos naturales.

La concepción cognoscitivista objetivista naturalista fue duramente criticada a principios del siglo xx por G.E. Moore (1903), mediante su famoso argumento de la *pregunta abierta*. Según el autor, la posición naturalista no podía ser correcta porque si los juicios morales fueran reducibles a enunciados descriptivos sobre propiedades naturales, no tendría sentido preguntarse si la propiedad o circunstancia  $x$  (que supuestamente determina su valoración moral) es ella misma buena, pues en tal caso afirmar por ejemplo «lo que produce placer el bueno» sería como la simple tautología «lo que produce placer produce placer», y sin embargo tiene sentido preguntarse si lo que

produce placer (o cualquier otra circunstancia o propiedad natural) es bueno. Ello implicaría, según Moore, que «bueno» como categoría moral, no puede consistir en ninguna propiedad natural o empírica. La propuesta de Moore, a la que a grandes rasgos se puede también añadir a autores como W.D. Ross (1930) o H.A. Prichard (1949), es que los conceptos morales básicos (como «bueno», «correcto» o «justo») describen propiedades no naturales o metafísicas, a cuyo conocimiento se accede a través de la intuición moral (de ahí que estas posiciones también suelen ser denominadas como *intuicionismo ético*).

Por su parte, dentro del cognoscitivismo subjetivista también es posible diferenciar entre distintas corrientes o concepciones en función de cuál es el sujeto respecto al cual se refieren las creencias y actitudes relevantes para determinar lo que se considera como moralmente correcto o incorrecto. Una de tales concepciones es la de la *teoría del mandato divino*, conforme a la cual un acto es moralmente bueno o correcto si se corresponde (o, al menos, resulta compatible) con lo que ordena la divinidad. Así, «*x* es moralmente bueno/correcto» significaría «ordenado o aprobado por Dios», mientras que «*y* es moralmente malo/incorrecto» significaría «prohibido o rechazado por Dios», respectivamente.<sup>18</sup>

Una concepción cognoscitivista subjetivista alternativa es la que a veces se denomina la *teoría del observador ideal*. En una gran simplificación, aquello moralmente correcto sería lo que determinaría como tal un sujeto hipotético con ciertas características ideales y en circunstancias óptimas, como ser totalmente racional, completamente imparcial y absolutamente informado de todos los hechos relevantes del caso.<sup>19</sup>

---

18 Puede hacerse aquí referencia a la clásica discusión medieval entre la posición *voluntarista* (como la de Guillermo de Ockham), según la cual *x* es bueno porque Dios así lo ordena, y la posición *intelectualista* (como la de Tomás de Aquino), conforme a la cual Dios ordena *x* porque *x* es bueno.

19 Esa sería por ejemplo la posición de W. Frankena (1963). Podría entenderse que esta posición es, al menos bajo cierta interpretación, la asumida por Kant mediante las exigencias de universalización del imperativo categórico, aunque es discutible que para el filósofo alemán los juicios morales fuesen descriptivos. También parecería ser, al menos bajo cierta interpretación (pues tampoco resulta evidente si concibe o no los juicios morales como enunciados descriptivos), la posición asumida por Carlos S. Nino en su modelo de constructivismo ético, en el que el autor argentino afirma explícitamente: «un juicio que expresa que moralmente debe hacerse *x* puede analizarse como un juicio que predica de la acción *x* que ella es requerida, en ciertas circunstancias definidas por propiedades fácticas de índole genérica, por un principio público que sería aceptado como justificación última y universal de acciones por cualquier persona que fuera plenamente racional, absolutamente imparcial y que conociera todos los hechos relevantes» (Nino, 1989:117 – cursivas en el original). Aunque muy interesante, esta concepción no parece dejar claro cuáles son los criterios sustantivos de corrección moral, pues la racionalidad, la imparcialidad y el conocimiento de los hechos no parecen ser tales criterios por sí mismos. Así, es razonable pensar que un individuo plenamente racional, imparcial e informado guiado por criterios utilitaristas llegará al menos en algunos casos a conclusiones distintas acerca de lo que resulta moralmente correcto que otra

Una tercera posición o conjunto de posiciones que puede incluirse en el cognoscitismo subjetivista es el de las concepciones relativistas, entendiendo por estas las que sostienen que lo moralmente (in)correcto depende de la opinión de cada persona o grupo en particular. Consiguientemente, afirmar «*x* es moralmente correcto» no significa, sino que cierto individuo o grupo social determinado aprueban *x*. En este sentido, proposiciones que en abstracto pueden ser contradictorias, como «el aborto es inmoral» o «el aborto está moralmente justificado al menos en algunos casos» pueden ser ambas verdaderas (o ambas falsas), en función del individuo o grupo al que se refieran. Otra consecuencia importante consiste en que, bajo esta concepción, al menos cuando las afirmaciones se refieren a un grupo social, desaparece la distinción entre el nivel de la ética normativa y el de la moral social o ética descriptiva, pues el primero colapsa en el segundo.

También puede considerarse como una concepción cognoscitivista, aunque un tanto sui generis, la llamada *teoría del error*, cuyo principal exponente es el filósofo australiano J. Mackie (1977). Esta curiosa y llamativa denominación responde a que los juicios éticos parecen tener una estructura descriptiva acerca del valor moral de las distintas acciones (por ejemplo, al sostener que «la tortura siempre es moralmente incorrecta» parece describirse alguna propiedad o característica moral objetiva de dicha acción), pero tales juicios son todos falsos porque no existen los valores morales objetivos, y por tanto se refieren a entidades inexistentes (de ahí el «error»). Ello conduce a los partidarios de esta teoría a adoptar una posición escéptica con respecto a la moral.

El otro gran conjunto de teorías metaéticas es de las concepciones no-cognoscitivistas, caracterizadas por sostener que la función primordial de los juicios éticos no es la de aportar información o describir algo, por lo que no son proposiciones. Históricamente, el primer conjunto de teorías metaéticas explícitamente no-cognoscitivistas fueron las *emotivistas* de filósofos como A. Ayer y Ch. Stevenson, si bien estas encuentran continuidad en tiempos más recientes a través de las concepciones *expresivistas* de autores como A. Gibbard (1990) o S. Blackburn (1993).<sup>20</sup> Para el emotivismo, un juicio moral como «la guerra nunca está moralmente justificada» no tiene por cometido describir nada (por más que pueda aportarnos información sobre las creencias y actitudes del sujeto), sino más bien *expresar* o *exteriorizar* (que no meramente informar o describir) las actitudes del sujeto frente a la guerra (en

---

persona también plenamente racional, imparcial e informada guiada por el imperativo categórico kantiano.

<sup>20</sup>En el caso de Blackburn, su posición también suele denominarse como *cuasirealismo*, al sostener que los juicios morales *expresan* o *proyectan* actitudes como si fueran propiedades reales u objetivas (aunque estrictamente no lo sean).

este caso, de rechazo), intentando con ello también influir en las actitudes de los demás (crear adhesión).

La otra corriente principal dentro de las posiciones no-cognoscitivistas es la representada por el *prescriptivismo*, cuyo defensor más destacado es el filósofo inglés R.M. Hare.<sup>21</sup> Según el autor, la función primordial del lenguaje moral es la guía de la conducta, es decir, dirigir el comportamiento de los individuos prescribiendo cómo tienen que comportarse, por lo que la estructura de los juicios morales se ajusta a la de las normas o prescripciones, y no a la de las proposiciones. Por tanto, una expresión como «lo moralmente correcto es cumplir con lo prometido» vendría a ser un modo de formular la norma o prescripción «deben cumplirse las promesas». Ahora bien, no cualquier prescripción es un candidato adecuado para ser un juicio moral, pues estos tienen que cumplir ciertas condiciones o exigencias, entre las cuales la más destacable es la de su universalización<sup>22</sup> (y por ello la posición de Hare suele también denominarse habitualmente como prescriptivismo universal).<sup>23</sup>

Este brevísimo repaso muestra que existen posiciones muy diversas en el campo de la metaética, y tal diversidad en principio facilitarí­a la aparición de desacuerdos en este ámbito. Nótese en cualquier caso que el hecho de que dos o más sujetos sostengan juicios morales incompatibles (esto es, acerca de qué es moralmente correcto o incorrecto), no implica que mantengan un desacuerdo en el ámbito metaético (aunque tampoco lo excluye), o viceversa. Como hemos ido viendo en los apartados anteriores, en lo que respecta a la presencia de desacuerdos, cada uno de los tres distintos niveles del discurso moral (ética descriptiva, ética normativa y metaética) es independiente de los demás, de modo que la (in)existencia de un desacuerdo en alguno de ellos no implica la (in)existencia en los demás. De ese modo, es posible tanto que dos sujetos que compartan una misma posición metaética sostengan juicios morales incompatibles, como que personas con concepciones metaéticas divergentes coincidan en sus juicios morales (además, claro está, de la posibilidad de que disientan o coincidan en ambos niveles).

Aclarado este punto, lo que sí puede afirmarse es que, en la medida en que cada una de las corrientes o concepciones mencionadas resulta, al menos en algún aspecto, incompatible con cualquiera de las demás, el mero hecho de

---

21 Ver, por ejemplo, Hare (1961).

22 Por esta razón, usando un ejemplo del propio autor, el racismo no estaría moralmente justificado, porque un nazi que afirmara «todos los judíos deben ser exterminados» debería estar dispuesto a aceptar que él mismo debiese ser exterminado en caso de ser judío, lo cual es como mínimo altamente improbable.

23 Bajo cierta interpretación, podría entenderse que esta es también la posición metaética asumida por Kant, en la medida en que los juicios morales son también concebidos como *imperativos* y que la exigencia de universalización es también un requisito fundamental de los mismos.

que distintos sujetos se muestren como defensores o partidarios de teorías metaéticas distintas conllevaría la existencia de un desacuerdo.

Pero incluso parece también existir espacio para los desacuerdos dentro del marco de una misma concepción metaética, al menos en algunas de ellas. Ese podría ser el caso, por ejemplo, de dos sujetos que, compartiendo una visión cognoscitivista objetivista naturalista de la ética, divergieran acerca de qué aspectos empíricamente observables constituyen el fundamento de los juicios morales.<sup>24</sup> Otras concepciones, que a primera vista no parecerían dejar espacio para los desacuerdos, podrían revelar en un examen más minucioso que sí que permiten tal posibilidad. Por ejemplo, dos partidarios de la teoría del mandato divino, pese a estar de acuerdo en que es moralmente correcto es aquello ordenado por Dios y coincidir en cuáles son las fuentes para determinar qué es aquello ordenado por Dios, pueden divergir respecto al *fundamento* de la corrección moral de lo que es ordenado por Dios (qué es lo que fundamenta o justifica que lo que Dios ordena es moralmente correcto, lo cual forma parte también de la reflexión metaética), como ilustra el tradicional debate entre las visiones *voluntarista* e *intelectualista*. Por otro lado, empero, parecen existir también concepciones metaéticas que no dejan espacio para la aparición de desacuerdos «internos», como en el caso de la teoría del error, para la que independientemente de las contradicciones entre los juicios morales resultantes, a nivel metaético no habría desacuerdo alguno por ser todos ellos falsos. Tampoco en las concepciones cognoscitivistas relativistas, pues aun cuando dos sujetos puedan realizar afirmaciones contradictorias acerca de los juicios éticos mayoritarios en una determinada comunidad, ambos coincidirían en que tales juicios consisten en proposiciones descriptivas acerca de las creencias y actitudes morales de dicho grupo social. De un modo similar, resulta cuestionable que exista espacio para el desacuerdo entre sujetos que compartan una misma concepción metaética no-cognoscitivista (como puede ser el emotivismo), por más que los juicios morales resulten incompatibles, en la medida en que todos ellos resultan de la manifestación o exteriorización de las actitudes de los sujetos.

¿Cómo se relaciona la existencia de desacuerdos en metaética con el concepto de error moral? Como ya se ha apuntado, una clase de desacuerdos morales en metaética consiste en que diversos sujetos sostengan teorías o concepciones alternativas, pues todas resultan en algún punto incompatibles. Pero tratándose de teorías filosóficas, para afirmar que al menos uno de ellos sostiene una concepción metaética «errónea», habría que partir del presump-

---

24 Por ejemplo, suponemos que tanto A como B comparten una visión utilitarista moral, pero A en una versión hedonista clásica, mientras que B en una versión centrada en la satisfacción de las preferencias informadas.

to de que contamos con criterios de corrección mínimamente claros y compartidos para valorar las teorías filosóficas como correctas o erróneas, lo cual resulta, cuanto menos, discutible. Valorar la «corrección» de una teoría filosófica no es como, por ejemplo, determinar la verdad o falsedad de una proposición empírica, y aunque ello no significa que no exista ningún tipo de criterio (como la coherencia interna, la compatibilidad con intuiciones básicas preanalíticas, la «resistencia» de la teoría a contraejemplos o a casos extremos, etc.), en la reflexión filosófica la solidez de las teorías depende fundamentalmente de la calidad de los argumentos esgrimidos en su defensa.

Pero por otra parte, lo que sí que ofrece cada concepción metaética (o la mayoría de ellas, al menos) son criterios para valorar la *corrección* o aceptabilidad de los juicios morales del nivel de la ética normativa, de manera que, internamente a la teoría, puede llegar a discriminarse entre juicios correctos y erróneos. Así, todas las teorías cognoscitivistas comparten la asunción básica de que los juicios morales expresan proposiciones asertivas (aunque de muy diverso tipo en función de la teoría concreta), por lo que para todas ellas habría un criterio fundamental de corrección: la *verdad*. De este modo, los juicios morales correctos o justificados serían los verdaderos, mientras que los erróneos serían los falsos. Qué sea aquello que describen las proposiciones condiciona también el modo en cómo puede determinarse (si es que es posible hacerlo) su verdad o falsedad. Así, para un cognoscitivista objetivista naturalista, la corrección o el error de un juicio moral dependerá del valor de verdad de proposiciones acerca de hechos, propiedades o circunstancias empíricamente observables. Para un cognoscitivista objetivista no-naturalista, de la verdad de proposiciones sobre propiedades no naturales o metafísicas (lo cual le ha generado muchas críticas por la imposibilidad de determinar la verdad de manera objetiva o intersubjetivamente, al no poder acudir a criterios distintos a las propias creencias). Las teorías cognoscitivistas subjetivistas del observador ideal, por su parte, deben afrontar tanto dificultades conceptuales (determinar con cierta precisión cuáles son las condiciones para la existencia de dicho observador), como epistémicas, pues por definición se trataría de situaciones contrafácticas, que nunca se dan en el mundo real. Los partidarios de la teoría del mandato divino parecen tener más fácil discernir entre juicios correctos y erróneos, al poder identificar las fuentes, aunque como concepción basada en la autoridad, de forma similar a los sistemas jurídicos, se enfrentan también a los problemas y dificultades propios de la interpretación de dichas fuentes. Las concepciones relativistas también basarían la corrección o el error en el valor de verdad de proposiciones empíricas, aunque esta vez referidas a las creencias y actitudes de un grupo social determinado. Por último, como ya se apuntó en su momento, para las teorías del error

no resulta relevante la distinción entre juicios morales correctos o erróneos, en la medida en que todos serían falsos (esto es, erróneos).

En lo que se refiere a las teorías metaéticas no-cognoscitivistas, al no concebir los juicios morales como proposiciones, los criterios de corrección no pueden basarse en la verdad, y de hecho, en las concepciones de tipo emotivista o expresivista, es incluso cuestionable que pueda hablarse propiamente de criterios de corrección, ya que los juicios consistirían en manifestaciones o exteriorizaciones de las actitudes y sentimientos del agente. En la medida en que estos juicios expresen realmente tales actitudes y sentimientos, no serían «erróneos», a pesar de que pudiera existir un desacuerdo moral entre distintos agentes. En contraste, las visiones prescriptivistas sí que cuentan con algunos criterios de corrección, aunque pueden variar en función del autor. Para R.M. Hare, por ejemplo, una condición indispensable es el de la universalización de la prescripción, pues no puede hablarse de un juicio moral genuino si el agente no está dispuesto a aceptar las consecuencias de dicha universalización. Esto, no obstante, no excluye la posibilidad de desacuerdos morales, y si los sujetos implicados asumen prescripciones incompatibles, pero están dispuestos a asumir su universalización, no podría decirse en propiedad que uno de los dos está equivocado. En una concepción como la kantiana (si se interpreta esta como una forma de no-cognoscitismo prescriptivista) sería más difícil que se dieran situaciones de juicios incompatibles que respetaran los criterios de corrección (pues estos incluyen también la exigencia sustantiva de tratar a todas las personas como fines en sí mismas), pero en el hipotético caso de que esto ocurriera, tampoco podría sostenerse que al menos uno de los sujetos está equivocado.

## **Recapitulación**

A partir de lo expuesto en los apartados anteriores, pueden señalarse de manera sintética las siguientes conclusiones:

- Por «desacuerdo moral» puede entenderse de manera genérica toda situación en la que dos o más sujetos sostienen sendas posiciones divergentes e incompatibles en alguna cuestión perteneciente al ámbito o discurso moral. Por «error moral» se entendería cualquier posición asumida por un sujeto en materia moral que se aparte de los criterios de corrección relevantes aplicables en cada situación o contexto.
- Dentro del discurso moral es posible diferenciar al menos entre tres niveles o dimensiones: la ética descriptiva o sociológica, la ética normativa, y la metaética. Los desacuerdos pueden manifestarse en cualquiera de los tres

- niveles, pero entre sí son totalmente independientes: ni la presencia de un desacuerdo en uno de ellos no implica que también exista desacuerdo en los demás, ni la inexistencia de un desacuerdo en uno de los niveles impide que puedan existir desacuerdos en el resto.
- En el ámbito de la ética descriptiva, un desacuerdo moral puede describirse como la situación en que dos o más agentes afirman sendas proposiciones acerca de las creencias y actitudes morales de un mismo grupo social de referencia que resultan mutuamente excluyentes (inconsistentes o contradictorias).
  - En el ámbito de la ética descriptiva, al estar compuesta por proposiciones sobre hechos observables relativos a las creencias y actitudes de los miembros de un grupo social, el error consiste en la falsedad de tales proposiciones. Si dos sujetos sostienen proposiciones mutuamente excluyentes sobre las creencias y actitudes morales de un mismo grupo social, al menos uno de ellos estará necesariamente en un error.
  - En el ámbito de la ética normativa, el desacuerdo se manifiesta cuando diversos sujetos mantienen juicios morales incompatibles. Partiendo del análisis del concepto de «obligación moral» de M.J. Zimmerman, la anterior afirmación puede concretarse en alguna (o ambas) de las situaciones siguientes: a) cuando diferentes sujetos o agentes mantienen sendas posiciones distintas e incompatibles (mutuamente excluyentes o contradictorias) acerca de la (in)existencia de un determinado mundo personalmente accesible en el que s hace o no hace A en T', y/o b) cuando distintos sujetos mantienen posiciones distintas y contradictorias acerca del valor deóntico asignado a un determinado mundo accesible en el que s hace o no hace A en T'.
  - La valoración de la corrección o del error de los juicios de la ética normativa necesariamente debe hacerse a partir de teorías que establezcan los criterios de justificación y corrección de tales juicios, lo que corresponde al ámbito de la metaética, a excepción de las proposiciones relativas al carácter personalmente opcional de una conducta, y por tanto acerca de la (in)existencia de un mundo accesible, para un sujeto s en un momento T'.
  - En el ámbito de la metaética, existe un desacuerdo moral entre diversos sujetos o agentes cuando estos sostienen posiciones distintas e incompatibles acerca de cualesquiera de las cuestiones que se consideran incluidas en dicho ámbito, como entre otras, la concepción semántica de los juicios morales, su fundamentación, la epistemología de la moral, o la ontología de los valores y principios morales. Partiendo de la clasificación de concepciones metaéticas expuesta en este trabajo, cuando diversos sujetos adopten concepciones distintas se planteará un desacuerdo moral, en la medida en que resultan mutuamente incompatibles, aunque cabe también la posibilidad



- de que, al menos en el marco de algunas de las concepciones expuestas, se plantee un desacuerdo entre sujetos cuyas posiciones metaéticas puedan ubicarse en una misma categoría general.
- Sobre el error moral en el ámbito de la metaética, es posible diferenciar dos niveles: a) Por una parte, el del análisis de la (in)corrección de los juicios de la ética normativa a partir de los criterios establecidos por cada teoría metaética en particular; y b) el de la valoración de la corrección de las propias teorías o concepciones metaéticas alternativas. En este último punto, dichas teorías pueden valorarse del mismo modo que cualquier otra teoría filosófica.

## Referencias bibliográficas

- Blackburn, S. (1993). *Essays in Quasi-Realism*. Oxford University Press.
- Brandt, R.B. (1959). *Ethical Theory*. Prentice May.
- Darwall, S.; Gibbard, A. y Railton, P. (Eds.) (1997). *Moral Discourse and Practice. Some Philosophical Approaches*. Oxford University Press.
- Frankena, W. (1963). *Ethics*. Prentice-Hall.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.
- Gowans, C.W. (Ed.) (1987). *Moral Dilemmas*. Oxford University Press.
- Mason, H.E. (Ed.) (1996). *Moral Dilemmas and Moral Theory*. Oxford University Press.
- Hare, R.M. (1961). *The Language of Morals*. Clarendon Press.
- Hudson W.D. (1970). *Modern Moral Philosophy*. MacMillan.
- Jørgensen, J. (1938). *Imperatives and Logic*. Erkenntnis, 7, 288–298.
- Mackie, J. (1977). *Ethics. Inventing Right and Wrong*. Penguin Books.
- Moore, G.E. (1903). *Principia Ethica*. Cambridge University Press.
- Nino, C.S. (1989). *Ética y derechos humanos*. Astrea.
- Prichard, H.A. (1949). *Moral Obligation*. Oxford University Press.
- Ross, W.D. (1930). *The Right and the Good*. Clarendon Press.
- Van Roojen, M. (2015). *Metaethics. A Contemporary Introduction*. Routledge.
- Von Wright, G.H. (1963). *Norm and Action. A Logical Enquiry*. Routledge and Kegan Paul.
- Zimmerman, M.J. (1996). *The Concept of Moral Obligation*. Cambridge University Press.

## 13. Metafísica y moral

Samuele Chilovi

Deberíamos reducir la contaminación del planeta en una medida mucho mayor que la actual. Los estados y ciudadanos que disponen de recursos económicos mayores tienen el deber (moral) de prestar ayuda a las personas que carecen del acceso a recursos vitales básicos. Nadie debería ser discriminado por su raza, identidad de género u orientación sexual. Causar el sufrimiento de seres sintientes sin una razón adecuada es moralmente incorrecto.<sup>1</sup>

Afirmaciones como estas parecen ser aptas para ser verdaderas o falsas, ya que aparentemente consisten en la atribución de alguna propiedad o relación (ser obligatorio, ser incorrecto, deber, etc.) a un conjunto de individuos, acciones, o estados de cosas. De hecho, no solo parecen ser aptas para la verdad o falsedad, sino que al menos algunas de ellas —u otras del mismo tipo— parecen ser verdaderas: es plausible pensar que algunas propiedades morales están instanciadas por algún objeto en algún momento.

Naturalmente, podríamos acabar descubriendo que, al fin y al cabo, no hay propiedades morales en el mundo, y que afirmaciones que parecen implicar su existencia o bien son sistemáticamente falsas, o bien consisten simplemente en la expresión de emociones u otros estados conativos. Y sin embargo, si dicha conclusión nos parece contraintuitiva, y aceptamos el compromiso implícito en nuestro discurso ordinario y realista, en un sentido mínimo, de que existen propiedades morales actualmente instanciadas,<sup>2</sup> surgen inmediatamente un gran número de preguntas.

¿Qué explica que algo sea bueno o malo, correcto o incorrecto, justo o injusto? ¿Son las propiedades (relaciones) evaluativas (ser bueno, malo, deseable, etc.) más fundamentales que las propiedades deónticas (deber, poder, ser (in)correcto, (in)justo, estar permitido, prohibido, etc.), o *viceversa*? ¿Hay propiedades morales fundamentales, o están todas instanciadas en virtud de propiedades no morales (naturales o descriptivas) más básicas? ¿Qué significa afirmar que los hechos morales se dan en virtud de hechos naturales, qué tipo de relación conecta lo normativo con lo descriptivo? ¿Están los hechos mora-

---

1 Para designar lo que en inglés se expresa con el término *right*, en este trabajo usaré en algunas ocasiones los términos «correcto» o «corrección», pero principalmente expresaré este concepto a través de los términos «obligación», «ser obligatorio» y «deber».

2 En línea con el uso filosófico convencional, «actual» y «actualmente» se usarán en un sentido modal, no temporal, para referirse a este mundo. De acuerdo con el léxico filosófico, el término «mundo» no se utilizará para denotar un planeta (por ej. el planeta Tierra), sino un universo entero.

les fundamentados en nuestras proyecciones, conceptos, acuerdos o deseos, o son, más bien, rasgos objetivos del mundo, existentes independientemente de nuestras creencias y aptitudes sobre ellos? ¿Pueden dos acciones ser iguales en todos sus aspectos descriptivos, y sin embargo diferir en su perfil normativo (al ser, por ejemplo, una correcta y otra incorrecta)? ¿Existen los principios morales? Si es que sí, ¿juegan algún papel en explicar la existencia de hechos morales particulares, o son, por lo contrario, nada más que generalizaciones sobre ellos?

Estas preguntas son, en un sentido amplio, preguntas metafísicas: tienen que ver con la naturaleza de una parte de la realidad, con la explicación (constitutiva, no causal) o reducción que de ella se puede dar en términos más básicos, o con el carácter y la estructura de la explicación de ciertas entidades. El propósito del presente trabajo es aclarar dichas preguntas, presentar algunas de las tesis que se han ofrecido en su respuesta, así como los problemas a los que se enfrentan, y los debates que han generado.

## Ética y metafísica

Uno de los objetivos centrales de la ética normativa es decirnos qué acciones o situaciones son buenas y malas, correctas e incorrectas, obligatorias y permisibles; qué debemos o tenemos razones para hacer o no hacer. Sin embargo, las teorías éticas no pretenden solo proporcionarnos una *lista* de acciones (in)correctas, de obligaciones y prohibiciones. Más bien, y más fundamentalmente, también tratan de proporcionar *explicaciones* de las propiedades normativas de las acciones. Su objetivo es decirnos *por qué* las acciones correctas son correctas, por qué tenemos las obligaciones que de hecho tenemos, y por qué las razones para que nos comportemos de cierta manera son las que son.

Pero, ¿qué tipo de explicaciones dan las teorías éticas, y en qué consisten esas explicaciones? Para arrojar algo de luz sobre estas cuestiones, comencemos por considerar algunos ejemplos.

Consideremos primero el debate en torno al consecuencialismo. En su forma más general, el consecuencialismo sostiene que las propiedades normativas de una acción dependen de sus consecuencias. Después, diversas tesis especifican esta visión general mediante una serie de características adicionales. Por ejemplo, especifican si las consecuencias que importan son solo las que afectan a los seres humanos o más bien a todos los seres sintientes; si las consecuencias relevantes para la evaluación de una acción son solo las que estén causadas por ella; especifican cuál es la naturaleza del valor (o los valores) o del bien cuya promoción se considera relevante, y la medida en la que tiene

que promoverse. En esta línea, por ejemplo, el consecuencialismo maximizador (*maximizing act-consequentialism*) sostiene que el hecho de que un acto sea obligatorio (o moralmente correcto) depende únicamente de si este produce al menos tanto bien general como cualquier otra alternativa.<sup>3</sup>

Pero, ¿cómo hay que entender la noción de dependencia que estas teorías involucran? En primer lugar, cabe señalar que el tipo de dependencia en cuestión no es de carácter causal. Por un lado, al decir que las acciones correctas (obligatorias) son correctas en virtud del hecho de que maximizan el bien, se está afirmando la existencia de una relación *sincrónica* entre las propiedades deónticas y evaluativas de un acto. Por lo tanto, no se está afirmando que el hecho de que un determinado acto maximizara el bien en un momento *t* *causó* que ese mismo acto fuera correcto en un momento posterior *t'*. Más bien, se está diciendo que el hecho de que un acto maximizara el bien en *t* hizo que ese mismo acto fuera correcto *en t*. Por otro lado, es común pensar que las propiedades éticas son causalmente inertes, es decir, incapaces de tener relaciones causales con otras entidades, y, *a fortiori*, incapaces de tener una relación causal con aquellas propiedades evaluativas con las que los consecuencialistas consideran que están relacionadas.

En segundo lugar, se podría intentar interpretar la tesis consecuencialista en términos modales, es decir, como afirmando que, necesariamente, una acción es correcta si y solo si (*syss*) maximiza el bien. Sin embargo, esta interpretación pasaría por alto un elemento clave de la postura en cuestión. Cuando un consecuencialista afirma que las acciones correctas son aquellas que maximizan el bien, no está simplemente diciendo que las acciones correctas y las que promueven el bien son necesariamente co-extensivas —es decir, tales que necesariamente, cada acción que maximiza el bien es correcta, y *viceversa*—. Más bien, también pretende decirnos *por qué* las acciones correctas son las que son, y en este sentido afirmar la *prioridad* del bien (una propiedad evaluativa) sobre lo correcto (una propiedad deóntica). De hecho, incluso un no-consecuencialista podría estar de acuerdo en que siempre que una acción es buena, es correcta, y *viceversa*. Y sin embargo, las dos partes aún estarían en desacuerdo sobre una cuestión crucial, a saber, cuál de estas características normativas es explicativamente más básica, cuál de ellas explica la otra.<sup>4</sup>

---

3 Algunos proponentes clásicos del consecuencialismo son los utilitaristas (Bentham, [1789]1961), (Mill, 1861) y (Sidgwick, 1871).

4 Para ser más precisos (ver Berker, 2018), algunos noconsecuencialistas también podrían estar de acuerdo en que el ser bueno tiene prioridad ontológica respecto al ser obligatorio, siempre que proporcionen una explicación del valor no consecuencialista. Dejando de lado estas complicaciones, el punto importante es que una característica crucial de la disputa se perdería al interpretar las afirmaciones clave en términos modales.

Esta direccionalidad explicativa también se perdería si interpretamos el consecuencialismo como una teoría sobre la identidad de propiedades o como una tesis analítica. Ya que si se entendiera como la afirmación de que la propiedad de ser obligatorio (o correcto) es *idéntica* a la propiedad de ser bueno (o de maximizar el bien), eso no implicaría —como debería— que la propiedad de ser bueno tiene prioridad explicativa respecto a la propiedad de ser obligatorio. De la misma manera, si la tesis consecuencialista fuera entendida como un bicondicional verdadero analítica o conceptualmente, tampoco se podría afirmar una relación asimétrica entre las entidades involucradas (los conceptos «bueno» y «correcto»).

Estas dos interpretaciones también se enfrentan a otros problemas. Por un lado, además de ser aparentemente deficiente, la interpretación en términos de identidad parece ser demasiado fuerte. Para entender el porqué, tomemos como ejemplo la variante consecuencialista encapsulada en una interpretación del principio de utilidad, según la cual las acciones obligatorias son todas aquellas que maximizan la felicidad. Al afirmar dicho principio, un utilitarista no quiere necesariamente identificar la propiedad deóntica de ser obligatorio con la propiedad natural de maximizar la felicidad, ya que eso equivaldría a asumir un compromiso metaético sustantivo y altamente no trivial.<sup>5</sup> Más bien, sea como fuere la manera en la que hay que articular exactamente dicho principio, parece un requisito mínimo que esta deje espacio para la posibilidad de que las propiedades normativas no sean propiedades naturales.

Las reducciones analíticas, por otro lado, se enfrentan a un conjunto distinto de desafíos. Por un lado, entender el principio de utilidad como una conexión analítica o conceptual entre el concepto «obligación» y el concepto «maximización de la felicidad» entraría en conflicto con el argumento de la pregunta abierta de Moore (Moore, 1903 – sección 13), que cuestiona precisamente la plausibilidad de este tipo de relaciones conceptuales o analíticas. Y de nuevo, aunque el argumento de Moore podría eventualmente resultar ser poco sólido, parecería irrazonable exigir que el utilitarismo (o cualquier otra teoría estructuralmente análoga) demuestre que Moore está equivocado para defender con éxito su punto de vista.

Además, un problema general para las interpretaciones metalingüísticas es que las explicaciones que proponen deben poder ser expresadas en términos tales que los usuarios competentes de los términos y conceptos involucrados sean capaces de entender (meramente en base a sus habilidades lingüísticas). Sin embargo, parece plausible que las explicaciones proporcionadas por las teorías éti-

---

5 Volveremos a este punto en la siguiente sección, donde examinaremos diversas interpretaciones de diversos puntos de vista metaéticos, así como de la relación entre puntos de vista éticos y metaéticos.

cas no tengan necesariamente que ser accesibles a cualquiera usuario competente del vocabulario moral, ni siquiera implícita o tácitamente.

Hemos visto por tanto que no es plausible concebir el debate sobre el consecuencialismo como si tuviera que ver con tesis modales, analíticas, causales, o de identidad. Otros debates en ética normativa pueden entenderse como desplegando una estructura similar. Veamos tres ejemplos más.

G.E. Moore y H. Sidgwick parecieron aceptar dos interpretaciones distintas del consecuencialismo (ver Moore, 1903; Sidgwick, 1871).<sup>6</sup> Por un lado, ambos trataron de explicar todas nuestras obligaciones particulares apelando a una obligación general de maximizar el bien.<sup>7</sup> Por lo tanto, ambos estaban comprometidos con la idea de que todas las obligaciones (excepto una) se pueden explicar a partir de una obligación más básica de maximizar el bien, y que ninguna otra obligación es más fundamental que esta. Sin embargo, divergían acerca de la cuestión en torno a qué (si es que hay algo) podría explicar esta obligación básica. Aquí, precisamente, se detiene la cadena explicativa de Sidgwick, quien no ofrece ninguna otra explicación. Moore (1903), en cambio, pensaba que el hecho de que una acción maximice el bien es parte de lo que hace que esta sea obligatoria, y para él, esto explica por qué existe una obligación fundamental de este tipo. En otras palabras, la razón por la cual maximizar el bien es obligatorio (o correcto) es que esto es parte de lo que la propiedad de ser obligatorio (o correcto) esencialmente es.

A continuación, consideremos la teoría ética de W.D. Ross (Ross, 1930). Según este enfoque, las acciones tienen las propiedades deónticas de ser *prima facie* correctas o incorrectas, y las tienen en virtud de un conjunto limitado de características no deónticas. Por ejemplo, bajo esta perspectiva, el hecho de romper una promesa hace que una acción sea *prima facie* incorrecta. Luego se puede decir que un acto es correcto o incorrecto *simpliciter* en un caso particular si su grado de corrección o incorrección *prima facie* supera al de las alternativas. De esta manera, las propiedades deónticas quedan explicadas por un grupo de características subyacentes que contribuyen, de una manera no mecánica, a determinarlas.

Finalmente, consideremos el debate sobre el lugar de las razones dentro del ámbito normativo. Los defensores de un enfoque de «las razones en primer lugar» (*reasons first*) sostienen que la noción de razón es normativamente básica: todos los hechos normativos se explican sobre la base de hechos que versan sobre razones, y ningún otro hecho normativo es más básico que los que se refieren a razones. Como consecuencia, por ejemplo, explican por qué

---

<sup>6</sup> A este propósito, véase Schroeder, 2017.

<sup>7</sup> Aunque, para ser precisos, Moore expresó esto en términos de corrección (*right*).

alguien tiene la obligación de realizar un determinado acto en términos del balance entre las razones a favor y en contra de realizar ese acto, y no proporcionan ninguna explicación (que apele a otras nociones normativas) de lo que es para un hecho ser una razón.<sup>8</sup>

Todos estos ejemplos parecen estar relacionados con cuestiones de prioridad y explicación metafísica (Berker, 2018). Tanto si nos preguntamos si promover el bien (o hacer algo bueno) tiene prioridad respecto al tener una obligación (o hacer algo correcto), o si existen obligaciones generales inexplicables, o si existen propiedades deónticas *prima facie* que se explican en la forma propuesta por Ross, como si nos preguntamos si las razones forman una categoría normativamente fundamental, parece que lo que nos interesa es el orden de prioridad ontológica de algunas nociones normativas respecto a otras, y el lugar que los hechos normativos ocupan en la cadena explicativa.

Además, interpretar estas posturas en términos modales, lingüísticos o causales, o como tesis de identidad, es problemático por todas las razones mencionadas anteriormente. Ello se debe a que no solo dichas interpretaciones no capturan la noción de dependencia, y las nociones afines de fundamentalidad y explicación que la acompañan, sino que también corren el riesgo de distorsionar estas posturas y los debates que versan sobre ellas.

En la medida en que las teorías éticas consiguen proporcionar una explicación de los hechos normativos de interés que sea a la vez precisa y esclarecedora, su objetivo se habrá logrado. Es decir, en la medida en que la explicación que proporcionen sea tanto extensionalmente correcta —en el sentido de que clasifique las acciones buenas como buenas, las permisibles como permisibles (y así sucesivamente)— como explicativamente adecuada —en el sentido de que representa fielmente la estructura explicativa subyacente— habrán alcanzado su meta.

Sin embargo, incluso cuando esto se logre, aún quedarán por abordar algunas cuestiones clave, en particular con respecto a cómo debemos interpretar las explicaciones que proporcionan estas teorías, así como la noción de dependencia que manejan. Porque, dependiendo de la interpretación que demos, estas teorías resultarán ser más (o menos) compatibles con diversos enfoques, tanto con respecto a la relación entre distintas características éticas, como con respecto a la relación entre características éticas y no éticas. Proporcionar una descripción de estas relaciones, y un mapa de las posturas que

---

<sup>8</sup> Sin embargo, los defensores del enfoque de «las razones en primer lugar» están en desacuerdo sobre la cuestión de si las razones son completamente fundamentales. Bajo una interpretación no-naturalista de este enfoque lo son Parfit (2011), Scanlon (2014); mientras que bajo una interpretación naturalista no lo son, ya que a su vez deben explicarse en términos no normativos, Schroeder (2014), Street (2006).

resultan de las diferentes interpretaciones, cae dentro del ámbito de la metaética, tal y como veremos a continuación.

## **Metaética y Metafísica**

Naturalismo y no-naturalismo ético: observaciones preliminares

Uno de los principales debates de la metaética contemporánea gira en torno a la cuestión de si las propiedades y relaciones morales son *naturales*. Aunque no es tarea fácil definir —casi nunca lo es— la noción de propiedad moral o natural, se puede dar una aproximación intuitiva por medio de ejemplos.

Como sabemos, rasgos morales paradigmáticos son los de ser bueno, correcto, obligatorio, permisible, etc. Por otro lado, la noción de propiedad natural relevante en este contexto es una noción muy inclusiva. Comprende no solo las propiedades estudiadas por las ciencias naturales (ser un electrón, tener una cierta masa, ser un organismo vivo, etc.), sino también características descriptivas no fundamentales presentes en el discurso ordinario (causar sufrimiento, romper una promesa, traicionar un amigo, etc.). Esto implica que en el sentido relevante, cualquiera característica descriptiva no normativa cuenta como natural en el contexto de este debate. (De esta manera, se puede definir la noción de hecho moral como un hecho que tiene al menos algún constituyente moral, y la de hecho natural como un hecho cuyos constituyentes son completamente no normativos).<sup>9</sup>

En una primera aproximación, el naturalismo ético defiende la tesis de que las propiedades y los hechos morales son naturales. El no-naturalismo, por el contrario, niega esta afirmación y sostiene que los hechos y propiedades morales son «distintos» o *sui generis*. Por supuesto, cualquiera puede ver que hay *alguna* diferencia entre hechos de este tipo —entre, digamos, el hecho de que comer carne de granjas industriales está mal, y el hecho de que las granjas industriales causan un gran daño a los animales que muchas personas comen—. La cuestión es si las diferencias son tan profundas como para excluir cualquier dependencia o reducción de los primeros hechos a los segundos, lo que significaría que las propiedades morales constituirían características genuinamente «novedosas» o «emergentes». En la metáfora tan utilizada, la pregunta es si, habiendo puesto en su lugar los hechos naturales, Dios terminó su trabajo (ya que todo lo demás se derivaría de ellos), o si por el contrario, faltaría que añá-

---

<sup>9</sup> En este trabajo, la noción de «hechos» que utilizo es una bajo la cual los hechos son estados de cosas, es decir, entidades del mundo compuestas por constituyentes como objetos, propiedades y relaciones.



diese los hechos morales básicos, sin los cuales la realidad moral no existiría.

Dejando de lado las metáforas, nuestra pregunta es ¿qué significa que una propiedad o relación moral *sea* natural?

Una propuesta inicialmente atractiva sería sostener que una propiedad cuenta como natural si existen condiciones no normativas que son modalmente necesarias y suficientes para su instanciación. Es decir:

–(Modal): Una propiedad  $F$  es natural si y solo si existe una condición no normativa  $\phi$  tal que necesariamente, para cada  $x$ :  $x$  es  $F$  si y solo  $x$  es  $\phi$ .

Dada esta caracterización de lo que hace que una propiedad sea natural, para que el naturalismo fuera cierto bastaría con que toda propiedad moral fuese necesariamente co-extensiva con alguna propiedad descriptiva.

Esta propuesta, aunque tradicionalmente haya sido popular, se enfrenta a dos tipos de problemas. El primero es que trivializa el debate a favor del naturalismo, al hacer que su verdad sea demasiado fácil de establecer. Esto se debe a que una de las tesis más generalmente aceptadas en metaética es la superveniencia de lo normativo en lo no normativo. Tanto los naturalistas como muchos no-naturalistas tienden a creer que si dos acciones (objetos, entidades) son iguales en todos los aspectos no normativos, entonces también deben de ser iguales en los aspectos normativos. O, dicho de otro modo, cada diferencia normativa entre dos entidades requiere alguna diferencia no normativa entre ellas:

–(Superveniencia): Para cualesquiera mundos  $v$  y  $w$ , y cualesquiera entidades  $x$  en  $v$  e  $y$  en  $w$ , si  $x$  en  $v$  e  $y$  en  $w$  son no-normativamente indiscernibles, entonces son normativamente indiscernibles.

Este principio nos dice que cada dos entidades (acciones, objetos, ...) posibles que tienen exactamente el mismo perfil no normativo —es decir, comparten todas sus propiedades no normativas, tanto intrínsecas como extrínsecas— también deben tener el mismo perfil normativo. Para tomar un ejemplo, supongamos que Víctor abandonó a su perro por razones fútiles y que su acto fue incorrecto. Entonces, cualquiera que en la misma situación —cualquiera que compartiendo todas las propiedades extrínsecas y relacionales de Víctor (sus motivos, sus intenciones, sus circunstancias, etc.)— hiciera lo mismo, estaría también haciendo algo incorrecto. Esto se debe a que, para casi todos los que se han ocupado de este tema, parecería arbitrario que dos acciones que son exactamente iguales y se realizan en las mismas circunstancias, pudieran ser tales que una de ellas sea correcta y la otra incorrecta, o una

buena y la otra mala.<sup>10</sup> En otras palabras, si fijamos los hechos no normativos, los hechos normativos deben ser los mismos. O, si consideramos dos acciones exactamente iguales en todas sus características no normativas, estas deben ser idénticas con respecto a su estatus normativo.

Sin embargo, el problema es que, dadas unas premisas plausibles, (Supervenencia) implica que cada propiedad moral es modalmente equivalente a alguna propiedad natural. Tomemos, por ejemplo, la propiedad de ser bueno y consideremos todas las acciones buenas, tanto las reales como las que son meramente posibles. Obviamente, cualquier buena acción (al ser una acción) también tendrá rasgos no normativos. Así que cada acción buena tendrá un perfil descriptivo completo —es decir, una propiedad compleja que codifica toda la verdad no normativa sobre ella—  $D_i(x)$ . Ahora consideremos la disyunción de los perfiles no normativos de todas las acciones buenas posibles:  $D_a(x) \vee D_b(x) \vee \dots$  y llamemos a la propiedad disyuntiva resultante  $\Delta(x)$ . Dada (Supervenencia), se deduce que necesariamente, para cada  $x$ :  $G(x) \leftrightarrow \Delta(x)$  — toda acción buena es  $\Delta$ , y toda acción que es  $\Delta$  es buena. De la misma manera, si repetimos este proceso para todas las propiedades morales, resultará que cada una de ellas es modalmente co-extensiva con una propiedad descriptiva —este argumento fue originalmente propuesto por Kim (1984); véase también Jackson (1998) y Rosen (2017a).

Esto significa que si (Modal) es cierto, el naturalismo estaría implicado por (Supervenencia). Y esto es problemático, ya que implicaría que quien rechaza el naturalismo también debe rechazar (Supervenencia), mientras que — pese a considerar las propiedades morales como *sui generis*— la mayoría de defensores del no-naturalismo todavía aceptan (Supervenencia).

Cabe señalar en este contexto que (Modal) es una consecuencia (casi irresistible) del «intensionalismo», es decir, del enfoque tradicional sobre la identidad de las propiedades, según el cual dos propiedades son idénticas si son necesariamente co-extensivas (ver Rosen, 2017a:153).<sup>11</sup> Por esta razón el argumento anterior, además de justificar el rechazo de (Modal), motiva también la búsqueda de un marco teórico en el que la equivalencia necesaria no sea suficiente para la identidad de propiedades. En otras palabras, un marco hipertextual.

---

10 Sin embargo, esta tesis ha sido cuestionada por algunos autores (véase especialmente Fine (2002), Hattiangadi (2018), Roberts (2017), Rosen (2020).

11 Supongamos que una propiedad  $F$  es equivalente a una propiedad no normativa  $G$  (necesariamente, para todo  $x$ :  $x$  es  $F$  si, y solo si,  $x$  es  $G$ ). Dado el intensionalismo,  $F$  y  $G$  son entonces propiedades idénticas. Y dado que  $G$  es natural,  $F$  también debe de serlo. A la inversa, supongamos que la propiedad  $F$  no es equivalente a la propiedad  $G$ : existen dos cosas que son ambas  $G$ , pero no ambas  $F$ . Entonces,  $F$  no sobreviene en  $G$ , y, por lo tanto, no es natural (ya que típicamente se supone que el naturalismo requiere supervenencia).

El segundo problema de (Modal) es similar al que afectó a las formulaciones modales de las teorías éticas. Aquí también, el problema es que la tesis que nos interesa entender (el naturalismo en este caso) trata de establecer una relación de *dependencia asimétrica* entre los hechos morales y naturales. Cuando los defensores del naturalismo sostienen que las propiedades y los hechos morales son naturales, lo que afirman es que los hechos morales se pueden derivar en función de los hechos naturales (y no *viceversa*). Dicho de otro modo, afirman que los hechos y propiedades naturales *determinan* los hechos y las propiedades morales, de tal forma que los hechos morales no podrían ser genuinamente «novedosos» o *sui generis*, al estar constituidos por hechos descriptivos.

Diversos autores han señalado que las relaciones modales de necesidad o superveniencia son incapaces de explicar de manera adecuada la noción de dependencia en cuestión. En este sentido, se han destacado dos problemas principales.<sup>12</sup>

El primer problema es que la superveniencia y la necesidad tienen las propiedades formales incorrectas para caracterizar la relación relevante de determinación, ya que ambas son reflexivas, no-simétricas (es decir, ni simétricas ni asimétricas), y monótonas.<sup>13</sup> En cambio, la noción de dependencia que, según la postura naturalista, se da entre propiedades morales y naturales, debería ser irreflexiva, asimétrica y no-monótona, ya que: (i) nada puede depender (en un sentido metafísico) de sí mismo; (ii) si  $x$  depende (metafísicamente) de  $y$ , entonces  $y$  no puede depender de  $x$ ; (iii) no es el caso que si  $x$  depende de  $y$ , entonces  $x$  depende de  $y$  junto con cualquier otra entidad  $z$ .<sup>14</sup>

El segundo problema deriva del hecho de que la superveniencia y la necesidad son relaciones intensionales, y por esta razón dan lugar a una serie de falsos positivos: casos de entidades entre las cuales se da una relación de su-

---

12 Véase especialmente Fine (1994), Rosen (2010), Schaffer (2009). Además, Berker (2018:9) hace notar, como Dancy (1981) y DePaul (1987), argumentos similares en contra de la identificación de la dependencia con la superveniencia, y en el contexto de la determinación de entidades normativas.

13 Una relación  $R$  es *reflexiva* *syss*:  $R$  se da entre cada cosa y sí misma —es decir, si para cada  $x$ ,  $Rxx$ ;  $R$  es *simétrica* *syss*: para cada  $x$  e  $y$ , si  $Rxy$ , entonces  $Ryx$ — es decir, si  $x$  está  $R$ -relacionado con  $y$ , entonces  $y$  está  $R$ -relacionado con  $x$ ;  $R$  es *asimétrica* *syss*: para cada  $x$  e  $y$ , si  $Rxy$ , entonces no es el caso que  $Ryx$ ;  $R$  es *monótona* *syss*: para cada  $x$ ,  $y$  e  $z$ , si  $x$  está  $R$ -relacionado con  $y$ , entonces  $x$  junto con  $z$  también está  $R$ -relacionado con  $y$ ;  $R$  es *no-monótona* *syss*: no es el caso que para cada  $x$ ,  $y$  e  $z$ , si  $x$  está  $R$ -relacionado con  $y$ , entonces  $x$  junto con  $z$  también está  $R$ -relacionado con  $y$ ;  $R$  es *transitiva* si para cada  $x$ ,  $y$  e  $z$ , si  $x$  está  $R$ -relacionado con  $y$ , y  $y$  está  $R$ -relacionado con  $z$ , entonces  $x$  está  $R$ -relacionado con  $z$ .

14 Este argumento ha sido defendido, entre otros, por Koslicki (2015:308), McLaughlin y Bennett (2021 – §3.5), Raven (2012:690), Raven (2013:194), y Schaffer (2009:364). Para un análisis más detenido de este argumento, véase Chilovi, 2022.

pervenencia y necesidad, pero que sin embargo no están ligadas por la relación de determinación. Por un lado, cada entidad necesaria sobreviene en, y está necesitada por, cualquier otra entidad (aunque claramente no se da, en muchos casos, una relación de dependencia entre ellas). Por otro lado, cada pareja de entidades necesariamente co-extensivas es tal que cada uno de sus miembros sobreviene en y está necesitado por el otro. Un caso de este tipo es el que utilizó K. Fine (1994) para argumentar en contra de las concepciones modales de la noción de esencia. El caso versa sobre la relación entre los conjuntos y sus elementos. Dado un conjunto determinado —{Sócrates}, digamos, es decir el conjunto cuyo único elemento es Sócrates— si ese conjunto no existiera, tampoco existiría Sócrates, y si Sócrates no existiera, tampoco existiría ese conjunto. Sin embargo —y este es el punto clave— la existencia del conjunto es la que depende de la existencia de sus elementos, no *viceversa*.

El caso de la relación entre las propiedades morales y naturales también es de este tipo. Como acabamos de ver, la supervenencia de lo normativo sobre lo no normativo, combinada con asunciones más que plausibles, implica que cada propiedad moral es necesariamente co-extensiva con alguna propiedad natural. Y sin embargo, la postura naturalista considera que existe una asimetría entre los dos conjuntos de propiedades, ya que son las propiedades morales las que dependen de las naturales, y no al revés.

## Fundamentación metafísica

Frente a estos problemas, varios autores han ido desarrollado una línea de investigación que, en lugar de tratar de analizar la dependencia constitutiva en términos modales o metalingüísticos, investiga dicha noción en sus propios términos. La teoría de la fundamentación (*metaphysical grounding*), desarrollada por primera vez en Bennett, (2011), Correia (2005, 2010), Fine (2001, 2012), Rosen (2010) y Schaffer (2009), es el resultado de este cambio de trayectoria.

A través de la noción de fundamentación se trata de capturar la relación de dependencia/determinación que se da entre entidades situadas en distintos niveles de la realidad, es decir, entre entidades más y menos básicas. Las entidades fundamentales son aquellas que no están fundamentadas en ninguna otra entidad, mientras que todas las demás —las que son «derivadas»— deben tener un fundamento completo en otras entidades más básicas. Por supuesto, qué entidades están vinculadas por una relación de dependencia metafísica, y qué entidades son fundamentales no son cuestiones triviales. Por ejemplo, según una concepción fisicalista, todas las entidades fundamentales son entidades físicas, y todo lo demás —los objetos y propiedades químicas,

biológicas, mentales, sociales, etc.— está, en última instancia, fundamentado en lo físico.<sup>15</sup> En cambio, bajo una postura dualista, algunas propiedades mentales serían básicas también.<sup>16</sup> (Otras entidades que han sido concebidas como posiblemente básicas son los objetos abstractos lógicos y matemáticos, y, lo que nos importa más aquí, algunas propiedades o principios morales).<sup>17</sup>

La teoría de la fundamentación desempeña su papel principal en la aplicación a debates filosóficos como los que acabamos de mencionar, ya que parte de su objetivo es precisamente permitirnos comprender mejor lo que está en juego en ellos. En este sentido, su utilidad derivaría del supuesto de que es la fundamentación la relación expresada por tesis filosóficas explicativas, al afirmar, por ejemplo, que un determinado enunciado es verdadero *porque* otro lo es, o que una determinada propiedad está instanciada *en virtud de* la instanciación de otra propiedad, en casos en los que las expresiones «porque» y «en virtud de» son utilizadas en un sentido constitutivo, no causal.<sup>18</sup>

Aunque queden muchas cuestiones abiertas respecto a la lógica y la naturaleza de la fundamentación, hay cierto consenso sobre algunos de sus rasgos más importantes. En primer lugar, cabe mencionar la distinción entre la fundamentación completa y la fundamentación parcial. Una pluralidad de hechos  $\Gamma$  fundamenta *completamente* algún hecho [p] si no es necesario agregar nada a  $\Gamma$  para tener una explicación completa de [p]. Correlativamente, [p] está *parcialmente* fundamentado en  $\Gamma$  si hay alguna pluralidad  $\Delta$  tal que [p] está completamente fundamentado en  $\Delta$ ,  $\Gamma$ . En segundo lugar, se suele suponer que la fundamentación es una relación fáctica, irreflexiva, asimétrica, transitiva y no-monótona, que se da (como mínimo) entre hechos.<sup>19</sup> Además, a diferencia de

---

15 Para una interpretación del fisicalismo como una tesis sobre la fundamentación, véase (Schaffer, 2017).

16 Para una defensa del dualismo, ver Chalmers (1996).

17 La noción de fundamentación ha sido recientemente utilizada para formular un amplio número de posiciones y debates filosóficos. Para una aplicación a la ontología social, véase Epstein (2015), Griffith (2018) y Schaffer (2019). Para una aplicación a la filosofía del derecho, véase Chilovi (2020) y Chilovi y Pavlakos (2019).

18 Para una ilustración de los casos que involucran la tesis de la fundamentación, véase especialmente Correia (2010), Rosen (2010) y Schaffer (2009, 2016).

19 Ver la nota 13 para una definición de estas propiedades formales. Audi (2012), Rosen (2010) y Schaffer (2009), por ejemplo, asumen la posición ortodoxa bajo la cual la fundamentación tiene esas características. Sin embargo, todas estas asunciones han sido cuestionadas. Jenkins (2011) y Kovacs (2018) plantean unos desafíos a la irreflexividad; para críticas a la asimetría, ver Thompson (2016); sobre la transitividad, ver Litland (2013), Loss (2017), Schaffer (2012) y Tahko (2013); sobre la monotonía, ver Audi (2012), Correia (2010) y Fine (2012), quienes proponen regimentar los enunciados de fundamentación no a través de un predicado relacional, sino a través de un conector. Schaffer (2009) defiende que la fundamentación es una relación que puede darse no solo entre hechos, sino entre entidades de cualquier tipo.

la superveniencia y de la necesidad, la fundamentación es una noción hiperintensional, es decir, puede no darse entre entidades co-intensionales.<sup>20</sup> Aun así, es común aceptar que sigue habiendo una relación muy estrecha entre la fundamentación y la necesidad, ya que si lo fundamentado (aun siendo distinto de sus fundamentos) no es *nada más que* sus fundamentos, entonces parece implausible que pueda variar independientemente de ellos. Por esta razón, se suele aceptar que si una pluralidad de hechos  $\Gamma$  fundamenta completamente [p], entonces es metafísicamente necesario que  $(\Lambda\Gamma \rightarrow p)$ : la existencia de lo fundamentado está implicada por sus fundamentos.<sup>21</sup>

## Naturalismo ético

Equipados con estos recursos conceptuales, podemos definir el naturalismo ético de la siguiente manera:<sup>22</sup>

(Fundamentación): Necesariamente, todos los hechos morales o bien son idénticos a, o bien están completamente fundamentados en, hechos no normativos.

Quien se detiene aquí y es capaz de demostrar que esta afirmación es correcta ya habría logrado reivindicar la idea central del naturalismo —la idea de que los hechos y las propiedades morales *son*, en el sentido relevante, naturales—. Por otro lado, esto deja abierta la posibilidad de una interpretación más fuerte. El naturalismo así definido implica que toda propiedad moral admite una explicación completa en términos naturales, donde la explicación podría diferir de un caso a otro, no solo con respecto a diferentes propiedades morales sino también con respecto a la misma propiedad moral. Por ejemplo, un acto puede ser incorrecto porque se incumplió una promesa, y otro acto puede ser incorrecto porque se realizó con la única intención de causar daño, sin que exista una única condición no normativa que explique uniformemen-

---

20 Más precisamente, la fundamentación es una noción hiperintensional en el sentido de que no es el caso que si  $x$  es un fundamento de  $y$ , y  $x$  es co-intensional con  $z$ , entonces  $z$  también es un fundamento de  $y$ ; y no es el caso que si  $x$  es un fundamento de  $y$ , e  $y$  es co-intensional con  $z$ , entonces  $x$  también es un fundamento de  $z$ .

21 Siguiendo a (Rosen, 2010), utilizo « $\Lambda\Gamma$ » para designar la conjunción de las proposiciones que corresponden a los hechos en  $\Gamma$ . Este principio ha sido criticado por Leuenberger (2014a) y Skiles (2015).

22 Ver Leary (2020) y Chilovi y Wodak (2022) para una formulación similar. Rosen (2017a) presenta una versión restringida de esta afirmación (ver la nota 26).

te por qué las acciones incorrectas son incorrectas. Esta visión equivaldría a una forma de naturalismo *no-reduccionista*.<sup>23</sup>

Sin embargo, un tipo diferente de naturalista podría no quedarse satisfecho o satisfecha con esto, queriendo afirmar la tesis más fuerte de que para cada propiedad moral  $M$ , existe una condición no normativa que explica *uniformemente* por qué  $M$  se ejemplifica en todos los casos. Esto significaría que para cada propiedad moral  $M$ , los hechos de la forma  $[M\alpha]$  se explican uniformemente por hechos que involucran una misma condición no normativa  $\phi$ . Esto se mantendría si, por ejemplo, los hechos acerca de lo que se debe hacer estuvieran todos completamente fundamentados en hechos sobre la maximización de la felicidad, y de manera similar para otras propiedades.

Un tipo de naturalismo aún más ambicioso puede apoyarse en la noción de esencia para hacer una afirmación aún más fuerte. El uso de esta noción se reanimó después del ataque a la concepción modal de esencia llevada a cabo por K. Fine (en Fine, 1994). Dicha concepción sostiene que el hecho de que una propiedad sea esencial para un objeto no es nada más que el hecho de que dicha propiedad sea necesariamente instanciada por ese objeto. Sin embargo, Fine argumentó que muchas propiedades están instanciadas necesariamente, a pesar de no formar parte de la esencia de los objetos que las instancian. (Volviendo al caso anterior —Fine, 1994—, aunque es necesario que Sócrates sea el miembro único de  $\{\text{Sócrates}\}$ , esto claramente no es parte de lo que Sócrates es, es decir, no es parte de su esencia). Esto le llevó a concluir que nuestra noción intuitiva de esencia es primitiva e inanalizable.

Según el enfoque neo-aristotélico propuesto por Fine, si bien sigue siendo cierto que si una propiedad  $F$  es esencial para un objeto  $x$ , entonces es metafísicamente necesario que si  $x$  existe, entonces  $x$  es  $F$ , no todas las propiedades que tiene un objeto necesariamente son esenciales para él. Desde este punto de vista, para cualquier entidad  $x$ , habrá una colección de proposiciones  $p_1 \dots p_n$  cuya verdad depende de la naturaleza de  $x$  y, en consecuencia, tales que, para cada  $p_i$  de la colección, es parte de la naturaleza de  $x$  que  $p_i$ . La esencia de  $x$  puede identificarse entonces con la clase de todas esas proposiciones, es decir, con la clase de las proposiciones que se encuentran en su naturaleza.

Haciendo uso de esta noción, se puede proponer una especie de naturalismo *reductivo*.<sup>24</sup> En esta forma de naturalismo, para cada propiedad moral hay alguna condición no normativa cuya instanciación explica uniformemente por qué algo tiene esa propiedad, y esto es parte de lo que esa propiedad es *esen-*

---

23 Esto es lo que Rosen (2017a) llama naturalismo no-reductivo de tipo 1.

24 Es una forma de naturalismo «reductivo» según la taxonomía esbozada por Rosen (2017a). Esto se suma, por supuesto, a la versión reductiva del naturalismo en la que las propiedades morales son idénticas a las naturales.

*cialmente*. Es decir, la condición no normativa que fundamenta sistemáticamente un determinado rasgo moral también figuraría dentro de su esencia, diciéndonos qué es ese rasgo. Para ilustrarlo, este tipo de postura naturalista podría decir que no solo cuando una acción es correcta es correcta en virtud de maximizar la felicidad, sino que esto también es parte de lo que la propiedad de ser correcto esencialmente *es*, es decir, parte de su identidad.

Llegados a este punto, vale la pena destacar que cualquier tipo de naturalismo enfrentará el siguiente problema. En el marco que estamos considerando, todos los naturalistas se comprometen a tener en cuenta cada hecho moral como completamente fundamentado en hechos naturales. Ahora bien, consideremos algún hecho moral, digamos,  $[Ma]$ , y su fundamento natural  $[Na]$ . El hecho de que  $[Na]$  fundamente completamente  $[Ma]$  también es un hecho. Y, dada la definición anterior de un hecho moral como aquel que involucra al menos un constituyente moral, esto también contará como un hecho moral. La pregunta, entonces, es si los naturalistas deben, y (de ser así) si *pueden*, explicar este tipo de hechos apelando únicamente a hechos no normativos.<sup>25</sup>

Se podría pensar que apelar a la noción de esencia resulte especialmente útil aquí. Esto se debe a que, de acuerdo con un enfoque destacado sobre lo que fundamenta los hechos sobre la fundamentación, estos están determinados por hechos sobre la esencia de lo fundamentado (ver Dasgupta, 2014; Rosen, 2010). Por ello, sería parte de la propiedad moral  $M$  que cualquier acción que sea  $N$  es  $M$ , y esto fundamentaría el hecho de que  $[Na]$  fundamenta  $[Ma]$ . Pero, nuevamente este hecho sobre la esencia de  $M$  contiene un constituyente moral. Y siguiendo el mismo razonamiento, este debería también explicarse en términos naturales. Llegados a punto, no está claro qué puede hacer un defensor del naturalismo.

Una posibilidad es que adopte la «teoría del colapso» (*collapse view*) defendida por Bennett (2011) y deRosset (2013), según la cual el hecho de que  $\Gamma$  fundamenta completamente  $[p]$  se fundamenta a su vez completamente en  $\Gamma$ . O podría estipular que los hechos sobre la fundamentación no cuentan como «morales» para los propósitos de este debate, aunque involucren elementos morales. También se podría argumentar que para poder reivindicar el naturalismo sea suficiente que los hechos morales *particulares* —los hechos sobre la instanciación de una propiedad o relación moral por un objeto particular— estén completamente fundamentados en hechos naturales.<sup>26</sup> Sin embargo, ninguna de estas estrategias parece del todo satisfactoria.

---

25 Ver Leary (2020) para una presentación de este problema, que es estructuralmente análogo al problema planteado por Dasgupta (2014) para el fisicalismo.

26 Esta, en efecto, es la estrategia desarrollada en Rosen (2017a), quien sostiene que el naturalismo se preocupa únicamente de buscar los fundamentos de hechos morales particulares. En este caso, la tesis naturalista resultante sería que cada hecho moral *particular* está completamente fundamentado en una colección de hechos descriptivos no normativos.



## No-naturalismo

Examinemos ahora a los oponentes del naturalismo. Dado que el no-naturalismo es la negación del naturalismo, como mínimo, está comprometido con que haya al menos algún hecho moral que no esté completamente fundamentado metafísicamente en hechos no-normativos. En consecuencia, dado el carácter exhaustivo de la distinción normativo/no-normativo, dada la suposición de que el ámbito moral es el ámbito normativo más básico, y excluyendo que pueda haber casos de fundamentación reflexiva o cadenas explicativas infinitas, se seguiría que al menos un hecho moral es fundamental (es decir, no fundamentado en nada). Y esto nos daría finalmente una definición satisfactoria de lo que significa que los rasgos morales sean genuinamente «novedosos» o *sui generis*.

Se pueden distinguir diferentes versiones del no-naturalismo dependiendo de cómo consideran que se fundamentan los hechos morales, si es que lo hacen de alguna forma. Una versión radical podría sostener que incluso los hechos morales *particulares* (por ejemplo, que un acto en particular es incorrecto) carecen de fundamento. Sin embargo, esta lectura (llamada «brutalismo» por Rosen, 2017) parece especialmente inverosímil, ya que al menos *esos* hechos morales deberían poder ser explicados de *alguna* forma. Por ello, la alternativa al brutalismo debe decir que si bien los hechos morales particulares siempre están completamente fundamentados, a veces carecen de un fundamento metafísico completo no normativo.

Hay dos formas principales de articular esta idea. Ambas coinciden en que los hechos morales particulares se fundamentan parcialmente en hechos naturales subyacentes, pero están en desacuerdo en cómo se fundamentan, y en si los principios morales también juegan un papel explicativo.

El no-naturalismo particularista niega la existencia de principios morales. Al evitar cualquier recurso a tales principios, y para evitar colapsar en el brutalismo, el no-naturalismo particularista sostiene que los hechos morales están completamente fundamentados en hechos naturales. Sin embargo, ¿cómo evita convertirse en una forma de naturalismo? La manera más plausible de alcanzarlo parece apelar a una especie de fundamento que no es metafísico; apelaría entonces a lo que Fine (2012) llama «fundamento normativo».

Para Fine, hay tres nociones explicativas diferentes de fundamentación, cada una de las cuales corresponde a una relación distinta. Además de la fundamentación metafísica, existiría una fundamentación de tipo natural, que se daría, por ejemplo, entre el hecho de que una partícula se está acelerando y el hecho de que alguna fuerza positiva actúa sobre ella; y existiría también una fundamentación de tipo normativo, que es la que conectaría hechos morales

y naturales (Fine, 2012:37). Según Fine, ninguna de estas relaciones de fundamentación es reducible a las demás, y cada una de ellas está conectada a un tipo diferente de modalidad —en concreto, a la necesidad metafísica, natural y normativa, respectivamente—. Entonces, para el no-naturalista particularista, los hechos morales se explican completamente por los hechos naturales subyacentes (ya que se fundamentan normativamente en ellos), pero a su vez son genuinamente novedosos o *sui generis*, ya que no están fundamentados metafísicamente.

Este enfoque se enfrenta a serios desafíos. Primero, además de generar una proliferación de relaciones primitivas,<sup>27</sup> esta visión implica que todos los hechos morales, incluso los hechos morales sobre acciones particulares, son metafísicamente básicos. En este sentido, los hechos morales particulares estarían a la par con los hechos físicos fundamentales, lo cual es algo que muchos encontrarían inverosímil (Rosen, 2017b).

En segundo lugar, este enfoque tiene problemas para dar cuenta de (Superveniencia). Como hemos mencionado anteriormente, es comúnmente aceptado que las propiedades normativas sobrevienen en las propiedades no normativas con necesidad metafísica. Ante esto, surge el desafío de cómo se puede explicar (Superveniencia). Esto se debe a que generalmente se considera que no debe haber conexiones necesarias brutas entre entidades distintas, es decir, ninguna conexión necesaria entre distintos existentes debe dejarse sin explicar.

La respuesta proporcionada por los defensores del naturalismo es simple: (Superveniencia) se mantiene porque las propiedades y los hechos normativos están fundamentados metafísicamente en hechos naturales, y porque la fundamentación metafísica implica la superveniencia de lo fundamentado en sus bases, con la fuerza de la necesidad metafísica.<sup>28</sup> Por el contrario, los brutalistas y los particularistas no-naturalistas no parecen tener nada que ofrecer a este respecto. En el caso del brutalismo, esto se debe a que no plantea ninguna conexión explicativa entre hechos morales y naturales; en el caso de los particularistas, se debe a que la conexión que postulan es de fundamentación normativa, lo que implica una forma de superveniencia más débil y que no llega a la necesidad metafísica.

La segunda forma en la que un no-naturalista puede explicar la relación entre hechos naturales y morales particulares es apelando a principios genera-

---

27 Para un argumento en contra de este tipo de pluralismo y en defensa de la unidad de la fundamentación, ver Berker (2018).

28 Para una crítica de la tesis de que la fundamentación implica superveniencia, ver Leuenberger (2014b); para una defensa de un principio de este tipo, ver Chilovi (2021).

les que los conectan.<sup>29</sup> Por ejemplo, diríamos que el acto de Víctor fue malo no solo porque rompió una promesa (un hecho natural particular), sino también porque romper las promesas está mal (un principio moral). En esta forma de naturalismo basado en principios, los hechos morales particulares se fundamentan metafísicamente en parte en hechos naturales subyacentes, y en parte en principios morales que proporcionan un vínculo entre las características morales y las naturales.

Cabe señalar que ni la existencia de tales principios, ni el papel que desempeñan en la fundamentación de hechos morales particulares, son rasgos necesariamente no-naturalistas. Los naturalistas también pueden apelar a estos principios al ofrecer explicaciones, siempre que los fundamenten en hechos descriptivos más básicos. Para los no-naturalistas, por el contrario, estos serán los candidatos naturales a ser los hechos morales fundamentales en la cadena explicativa (razón por la cual, aunque permiten que los hechos morales particulares estén completamente fundamentados, todavía niegan que estén totalmente fundamentados en hechos naturales).

Esta postura tampoco está exenta de problemas (véase Berker, 2019). El desafío principal radica en cómo deben entenderse los principios morales, y en particular en el hecho de que ninguna de las hipótesis principales sobre cómo concebir su forma y contenido nos permite otorgarle el rol explicativo que según esta postura juegan.<sup>30</sup>

Una posibilidad es que los principios sean entendidos como generalizaciones condicionales (o bicondicionales) que conectan propiedades morales y naturales. Por ejemplo, dada una propiedad moral *M* y una natural *N*, podría ser un principio moral que, para todo *x*: *x* es *M* si (y solo si) *x* es *N*. Sin embargo, si este fuera su forma, los principios no podrían explicar sus instancias (por ejemplo, el hecho particular [*Ma*]), porque, al ser generalizaciones, estarían ellos mismos fundamentados en sus instancias (y recordemos que la fundamentación es asimétrica).<sup>31</sup>

Un remedio natural sería entonces formular los principios como (bi)condicionales metafísicamente necesarios. Pero, si fuera así, surgiría entonces un dilema a partir del hecho de que o bien los principios así entendidos son fundamentales, o bien no lo son. Si son fundamentales, entonces se volvería a postular el mismo tipo de necesidades brutas que —como en el caso de (Superveniencia)— muchos consideran problemáticas. Si por el otro lado no

---

29 Esta es la posición que Rosen (2017a) denomina «no-naturalismo de la ley puente» (*bridge law non-naturalism*). Sobre esta posición, y sobre la manera en la que está conectada con la fundamentación normativa, véase también Moreso (2018).

30 Sobre esta cuestión, ver especialmente Berker (2019), Enoch (2019) y Rosen (2017a, 2017b).

31 Véase Rosen (2010).

lo fueran, estarían plausiblemente fundamentados en la esencia de las propiedades morales a las que hacen referencia.<sup>32</sup> Por lo cual, se seguiría de su verdad que hay condiciones naturales suficientes para la instanciación de propiedades morales, y que forman parte de la esencia de estas propiedades. Y esta afirmación, aunque no sea estrictamente equivalente al naturalismo reduccionista, se acerca mucho a él.

Otra posibilidad sería formular los principios con un contenido explicativo, por ejemplo, de la siguiente manera: necesariamente, para todo  $x$ , si  $Nx$  (o  $Mx$ ), entonces  $[Nx]$  fundamenta  $[Mx]$ . Sin embargo, la forma resultante de no-naturalismo no parece ser del todo estable o sólida. Ya que, en primer lugar, la verdad de estos principios parecería implicar que el naturalismo es cierto (al ser los fundamentos a los que se refieren exclusivamente naturales). Y en segundo lugar, porque —por la misma razón— los principios así interpretados implicarían que ellos mismos son explicativamente redundantes (al referirse su contenido exclusivamente a hechos naturales). En conclusión, no queda claro que estos problemas puedan solucionarse.<sup>33</sup>

### **Conclusión: la relación entre ética y metaética**

Comenzamos este trabajo con la idea de que las teorías éticas normativas tratan de proporcionar explicaciones de varias propiedades normativas, pero no estaba claro qué significaba eso exactamente (aunque pudimos ver algunas de las cosas que no significaba). Más adelante, pasamos a examinar la tesis naturalista de que las propiedades y los hechos morales son naturales, así como su negación por parte del no-naturalismo.

A la luz de esa discusión, ahora podemos ver cuáles son las opciones a partir de las cuales se pueden reconstruir diferentes interpretaciones del carácter explicativo de las teorías éticas. Dejando de lado el brutalismo, que niega lo que todas estas teorías parecen presuponer —es decir, que los hechos morales particulares dependen de hechos más fundamentales— todas las demás opciones quedan abiertas.

Un resultado de este análisis es que puede ayudarnos a aclarar la forma en que la ética y la metaética son entre sí independientes (como se suele sostener), pero también la forma en la que *no* lo son.<sup>34</sup> Son independientes en el sentido

---

32 Véase Fine (2002) y Rosen (2017a).

33 Si los principios no pueden jugar un rol explicativo, serían más bien análogos a las generalizaciones en las que consisten las leyes de la naturaleza según una perspectiva humeana.

34 Al menos cuando se trata de la conexión entre la ética normativa y el debate entre naturalismo y no-naturalismo.

de que sea cual sea la teoría metaética correcta, el resultado de las disputas morales sustantivas no se verá afectado. No influirá, por ejemplo, en la cuestión sobre si las posturas éticas correctas son las consecuencialistas o más bien las deontológicas. No lo hará, al menos, cuando estas se consideren al margen de la pregunta sobre cómo entender la noción de dependencia que utilizan.

Por otro lado, las cuestiones metaéticas que hemos analizado tendrán un impacto en la ética normativa, en la medida en que para que estas teorías sean del todo precisas, hará falta elegir entre distintas opciones metaéticas. En este sentido, tendrá que haber un único marco metaético que dé cuenta correctamente de la naturaleza de la dependencia normativa que involucran, así como de la estructura de las explicaciones que proporcionan. Y esto, a su vez, determinará qué teoría ética, entendida en su totalidad, es realmente verdadera.

En conclusión, y de manera quizás inesperada, algunos conceptos metafísicos pueden ser útiles para la filosofía moral, al ayudarnos a comprender mejor no solo cuestiones metaéticas, sino también las teorías éticas de primer orden y los temas que les preocupan.<sup>35</sup>

## Referencias bibliográficas

- Audi, P. (2012). Grounding: Toward a Theory of the In-Virtue-Of Relation. *The Journal of Philosophy*, (109), 685–711.
- Bentham, J. (1789/1961). *An Introduction to the Principles of Morals and Legislation*. Doubleday.
- Berker, S. (2018). The Unity of Grounding. *Mind*, 507(127), 729–777.
- Berker, S. (2019). The Explanatory Ambitions of Moral Principles. *Noûs*, 4(53), 904–936.
- Chalmers, D.J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chilovi, S. (2020). Grounding-based formulations of legal positivism. *Philosophical Studies*, 11(177), 3283–3302.
- Chilovi, S. (2021). Grounding entails supervenience. *Synthese*, 6(198), 1317–1334.
- Chilovi, S. (2022). Introducción a la Metafísica. En González, D. Lagier y Lariguet, G., *Filosofía para Juristas. Una Introducción* (pp. 79-112). Trotta.
- Chilovi, S. y Pavlakos, G. (2019). Law-Determination as Grounding: A Common Grounding Framework for Jurisprudence. *Legal Theory*, 1(25), 53–76.
- Chilovi, S. y Wodak, D. (2022). On the (in)significance of Hume's Law. *Philosophical Studies*, 179(2), 633–653.
- Correia, F. (2005). *Existential Dependence and Cognate Notions*. Philosophia Verlag.
- Correia, F. (2010). Grounding and Truth-Functions. *Logique et Analyse*, (53), 251–279.

---

35 Quiero agradecer especialmente a Gloria Andrada por sus comentarios y ayuda con la revisión lingüística de este texto. Gracias también a José Juan Moreso por sus comentarios a una versión previa de este texto.

- Correia, F., y Schnieder B. (Eds.). (2012). *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press.
- Dancy, J. (1981). On Moral Properties. *Mind*, (90), 367–385.
- Dasgupta, S. (2014). The Possibility of Physicalism. *The Journal of Philosophy*, (111), 557–592.
- DePaul, M.R. (1987). Supervenience and Moral Dependence. *Philosophical Studies*, (51), 425–39.
- Enoch, D. (2019). How Principles Ground. *Oxford Studies in the Philosophy of Mind*, 14, 1–22.
- Epstein, B. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford University Press.
- Fine, K. (1994). Essence and Modality. *Philosophical Perspectives*, (8), 1–16.
- Fine, K. (2001). The Question of Realism. *Philosophers' Imprint*, 1(1), 1–30.
- Fine, K. (2002). Varieties of Necessity. En Gendler, T. y Hawthorne, J. (Eds.), *Conceivability and Possibility* (pp. 253–282). Oxford University Press.
- Fine, K. (2012). Guide to Ground. En Correia, F. y Schnieder, B. (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 37–80). Cambridge University Press.
- Griffith, A. (2018). Social Construction and Grounding. *Philosophy and Phenomenological Research*, 2(97), 393–409.
- Hare, R.M. (1952). *The Language of Morals*. Oxford University Press.
- Hattiangadi, A. (2018). Moral Supervenience. *Canadian Journal of Philosophy*, (48), 592–615.
- Jackson, F. (1998). *From Metaphysics to Ethics*. Oxford University Press.
- Jenkins, C. (2011). Is Metaphysical Dependence Irreflexive? *The Monist*, (94), 267–276.
- Kim, J. (1984). Concepts of Supervenience. *Philosophy and Phenomenological Research*, (45), 153–176.
- Koslicki, K. (2015). The Coarse-Grainedness of Grounding. *Oxford Studies in Metaphysics*, 9, 306–344.
- Kovacs, D.M. (2018). What is Wrong with Self-Grounding? *Erkenntnis*, (83), 1157–1180.
- Leary, S. (2020). Normativity. Raven, M. (Ed.), *Routledge Handbook of Metaphysical Grounding* (pp. 472–483). Routledge.
- Leuenberger, S. (2014a). Grounding and Necessity. *Inquiry*, (57), 151–174.
- Leuenberger, S. (2014b). From Grounding to Supervenience? *Erkenntnis*, (79), 227–240.
- Litland, J.E. (2013). On Some Counterexamples to the Transitivity of Grounding. *Essays in Philosophy*, (14), 19–32.
- Loss, R. (2017). Grounding, Contingency and Transitivity. *Ratio*, (30), 1–14.
- McLaughlin, B. y Bennett, K. (2021). Supervenience. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition). <https://plato.stanford.edu/archives/win2018/entries/supervenience/>
- McPherson, T. y Plunkett, D. (Eds.). (2017). *The Routledge Handbook of Metaethics*. Routledge.
- Mill, J.S. ([1861]1998). *Utilitarianism*. Oxford University Press.
- Moore, G.E. (1903). *Principia Ethica*. Cambridge University Press.
- Moreso, J.J. (2018). Realismo Moral. En *Enciclopedia de la Sociedad Española de Filosofía Analítica*. <http://www.sefaweb.es/realismo-moral>.

- Parfit, D. (2011). *On What Matters*. Oxford University Press.
- Raven, M.J. (2012). In Defence of Ground. *Australasian Journal of Philosophy*, (90), 687–701.
- Raven, M.J. (2013). Is Ground a Strict Partial Order? *American Philosophical Quarterly*, (50), 191–199.
- Roberts, D. (2017). Why Believe in Normative Supervenience? En Shafer-Landau, R. (Ed.), *Oxford Studies in Metaethics*, (pp. 1–24). Oxford University Press.
- Rosen, G. (2010). Metaphysical Dependence: Grounding and Reduction. En Hale, B. y Hoffmann, A. (Eds.), *Modality: Metaphysics, Logic, and Epistemology* (pp. 109–135). Oxford University Press.
- Rosen, G. (2017a). Metaphysical Relations in Metaethics. En McPherson, T. y Plunkett, D. (Eds.), *Routledge Handbook of Metaethics* (pp. 109–136). Routledge.
- Rosen, G. (2017b). What is a Moral Law? *Oxford Studies in Metaethics*, 12, 135–159.
- Rosen, G. (2020). Normative Necessity. En Dumitru, M. (Ed.), *Metaphysics, Meaning and Modality: Themes from Kit Fine* (pp. 205–233). Oxford University Press.
- Ross, W.D. (1930). *The Right and the Good*. Oxford University Press.
- Scanlon, T.M. (2014). *Being Realistic about Reasons*. Oxford University Press.
- Schaffer, J. (2009). On What Grounds What. En Chalmers, D.J.; Wasserman, R. y Manley, D. (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology* (pp. 347–383). Oxford University Press.
- Schaffer, J. (2012). Grounding, transitivity, and contrastivity. En Correia, F. y Schnieder, B. (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 122–138). Cambridge University Press.
- Schaffer, J. (2016). Grounding in the Image of Causation. *Philosophical Studies*, (173), 49–100.
- Schaffer, J. (2019). Anchoring as Grounding: On Epstein's The Ant Trap. *Philosophy and Phenomenological Research*, (99), 749–767.
- Schroeder, M. (2014). *Explaining the Reasons We Share*. Oxford University Press.
- Schroeder, M. (2017). Normative Ethics and Metaethics. En McPherson, T. y Plunkett, D. (Eds.), *Routledge Handbook of Metaethics* (pp. 674–686). Routledge.
- Sidgwick, H. (1871). *The Methods of Ethics*. Thoemmes Press.
- Skiles, A. (2015). Against Grounding Necessitarianism. *Erkenntnis*, (80), 717–751.
- Street, S. (2006). A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies*, (127), 109–166.
- Tahko, T.E. (2013). Truth-Grounding and Transitivity. *Thought: A Journal of Philosophy*, (2), 332–340.
- Thompson, N. (2016). Metaphysical Interdependence. Jago, M. (Ed.), *Reality making* (pp. 38–56). Oxford University Press.

## 14. Agencia moral, libertad de voluntad y determinismo

*Fernanda Flores*

El presupuesto fundamental de toda ética normativa es que los agentes morales gozamos de libre albedrío y somos responsables de nuestra conducta. Así, en el dominio de la moral, la agencia aparece asociada a conceptos tales como el merecimiento, la culpa, el castigo y la autonomía, entre otros. Esta red conceptual expresa ideas fundamentales acerca de nuestros comportamientos y reacciones cotidianas, y también acerca de nuestro sistema de premios y castigos. La noción más extendida de responsabilidad moral remite a la atribución o merecimiento de aprobación o culpa, según se trate de un acto correcto o incorrecto, respectivamente. Dado que la libre voluntad suele ser considerada una condición necesaria para la agencia moralmente responsable, es plausible asumir que las personas en general poseen libertad de voluntad.

Sin embargo, al mismo tiempo, somos conscientes de que ciertos condicionamientos e influencias funcionan como atenuantes de la carga de responsabilidad que atribuimos a las personas y sus acciones. Por ejemplo, ante un hecho reprochable, no juzgamos del mismo modo a quien consideramos que actuó bajo condiciones psicológicas normales que a quien creemos que padece algún tipo de compulsión (adicciones, fobias, patologías, etc.) o a quien creemos que atravesó alguna situación extrema en su vida que lo llevó a actuar de ese modo. Estas actitudes a la hora de evaluar las acciones reflejan la idea de que los agentes no siempre son plenamente responsables de sus decisiones. En los actos inciden factores internos (patologías, habilidades, carácter) y externos (contexto, oportunidades) que escapan al control del individuo. Teniendo en cuenta estos factores, distintas disciplinas se esfuerzan por explicar la conducta de los agentes y esta explicación, además de valor epistémico, puede tener peso para la justificación de la acción, no solo para eximir moralmente a los agentes sino incluso, por ejemplo, para rebajar una pena.

A primera vista podemos notar cierta contradicción entre la afirmación de la libertad y la atención a los condicionamientos. Esta tensión puede pasar desapercibida en el orden cotidiano pero cobra especial relevancia en contextos filosóficos, ya que encubre una serie de problemas que ha generado interés en subáreas como filosofía de la acción, metafísica, gnoseología, ética y metaética. El análisis de las principales cuestiones que atañen a esta disciplina será expuesto en primer lugar, atendiendo a los conceptos básicos y cuestiones preliminares relacionadas, seguido del desarrollo de las principales corrientes a partir de la división tradicional compatibilismo–incompatibilismo,



y finalizando con una mención a algunos de los autores y autoras referentes de la temática en el último tiempo.

### **Conceptos básicos, niveles de análisis y tensiones: algunas cuestiones preliminares**

¿Qué entendemos por libertad de voluntad o de acción?<sup>1</sup> Claramente no hay una respuesta unánime. En la vasta bibliografía clásica y contemporánea sobre el tema, podemos hallar distintas respuestas y tematizaciones de los problemas, no siempre vinculados especialmente al aspecto ético o práctico. Aquí me interesa tratar la cuestión de un modo que sea relevante para pensar la responsabilidad moral. En este sentido, una de las concepciones más extendidas define la libre voluntad como la capacidad para ejercer *control* sobre nuestras acciones (McKenna y Pereboom, 2016; Haji, 2009; Mele, 2006). Esto implica la capacidad de sortear impedimentos externos e internos, y el control debe ser lo suficientemente relevante (tener determinado grado) en un acto como para atribuirlo al agente (Mele, 2006). En otras palabras, las acciones deben poder ser —al menos en alguna medida— dependientes del agente. Una idea similar ya se encontraba en la concepción de las acciones voluntarias de Aristóteles, siendo estas las que tienen como principio al individuo y este conoce las circunstancias de la acción, en contraste con aquellas en las que actúa por la fuerza o por ignorancia.<sup>2</sup> En principio, es plausible la idea general de que, por ejemplo, si alguien provoca un accidente a causa de un movimiento involuntario, no diríamos que es moralmente responsable, puesto que precisamente no se hallaba en control de su cuerpo.<sup>3</sup>

Otra forma de concebir la libertad de voluntad, extendida en el campo de la metaética (por ej., Moore, 1912; Ayer, 1954; Smith, 2003), la comprende

---

1 Varios autores (por. ej., Haji, 2009; Mele, 2006; Van Inwagen, 1983) ponen en duda la existencia de la voluntad como facultad y prefieren hablar en términos de *acciones libres* para no comprometerse con la existencia de poderes mentales causales. No obstante el sentido que le atribuimos, asociado al control, es lo suficientemente amplio para discutir el núcleo de problemas que refiere a la libertad de voluntad y la libertad referida a las acciones (sigo en este abordaje a McKenna y Pereboom, 2016:9–11).

2 Cfr. EN III i 1111a22. Sin embargo, tengamos presente que en la cosmovisión griega no había discusión libertad–determinismo en el sentido en que empezó a comprenderse a partir de los desarrollos científicos del siglo XVII.

3 Asimismo Talbert (2019) realiza una distinción útil entre responsabilidad causal y moral. Un agente, volviendo al caso del accidente, puede ser el responsable causal pero no moral. Además de este ejemplo podemos pensar el caso de los niños, quienes cuentan con aptitudes que les permiten llevar a cabo acciones y ser responsables causales pero no cumplen con otras condiciones para ser tratados como agentes moralmente responsables.

como *alternativas posibles*, también formulada como la capacidad o habilidad de hacer las cosas de otro modo. La idea básica es que para afirmar que un agente es libre este debe contar con alternativas de cursos de acción. Retomando lo anterior, la acción libre involucraría entonces dos capacidades o habilidades: la de ejercer el control y la de poder hacer las cosas de otro modo. Para muchos autores ambas se hallan interrelacionadas, para algunos la primera requiere la segunda,<sup>4</sup> y otros, como veremos con Frankfurt, niegan que la existencia de alternativas sea una condición necesaria para atribuir responsabilidad al agente.

Por otro lado, así como resulta complejo definir la libertad, sucede lo mismo con la caracterización de nuestra voluntad como determinada, sometida a una ley o leyes universales. En su forma más general, el determinismo presenta dos elementos básicos: uno temporal y otro de relación causal. El primero se expresa en la tesis que sostiene que en un momento dado hay solo un futuro físicamente posible (Mele, 2006:3; Van Inwagen, 1983:3), y el segundo expresa la idea de que todo evento está determinado por uno que le antecede (Hofer, 2016).<sup>5</sup> El determinismo se discute generalmente en los planos físico y metafísico, aunque también debemos agregar un plano teológico que tuvo su auge en el Medioevo y continuó en la Modernidad. Sostener la verdad del determinismo en términos metafísicos no supone la creencia de que el determinismo es verdadero en el mundo físico. Además, con relación al aspecto práctico de estas tesis es importante destacar que aun si se considera que el determinismo físico es verdadero, esto no implica que debemos adoptar una visión escéptica respecto a la moral —como veremos en la propuesta de Pereboom—. Esta distinción es pertinente porque muchas veces los filósofos son ambiguos en relación con la distinción entre niveles fáctico y moral, y la compatibilidad o incompatibilidad puede darse en cualquiera de estos órdenes o también en un nivel analítico (cfr. Double, 2002:508).

Otra cuestión problemática es que la libertad de la voluntad no solo parece ser irreconciliable con el determinismo sino también con el indeterminismo, que es la tesis que afirma que hay más de un futuro posible y esto tiene un correlato a nivel cuántico. Un tipo de indeterminismo que debemos destacar es el naturalismo. De acuerdo con esta teoría, la agencia humana se configura a partir de estados, eventos y procesos que están regidos por leyes naturales. Así toda acción humana puede explicarse causalmente, preferentemente en

---

4 Este requerimiento puede ser sustantivo (McKenna y Pereboom, 2016; Kane, 1996) o formal (Ginet, 1990; Van Inwagen, 1983).

5 En cuanto a la noción de leyes o ley de la naturaleza, implicada en este debate, basta con señalar que refieren a regularidades relativas al despliegue de los fenómenos en el mundo que implican necesidad, no son accidentales y establecen límites.

términos de probabilidad. Además —aunque suene poco intuitivo— determinismo e indeterminismo son visiones que pueden llegar a ser complementarias, ya que en un macronivel (el nivel de los fenómenos físicos, en oposición al nivel cuántico) manteniendo un enfoque indeterminista las relaciones causales pueden interpretarse como determinadas.

Por último podemos destacar que es posible distinguir dos grandes tendencias, referidas a la evidencia y la fuente de conocimiento que tomamos como válida. Por un lado, hay una tendencia *subjetivista*, basada en la percepción de la primera persona, en la evidencia de la experiencia propia y la toma de decisiones. Los individuos nos percibimos como libres y como seres capaces de deliberar y de ejercer autocontrol. Esta idea suele ir acompañada de la creencia de que somos libres de hecho, es decir, adoptamos una posición *realista* respecto al ejercicio de la libertad. Y por el otro lado, hay una tendencia *objetivista*, continuadora de la «imagen científica del mundo», que busca subsumir los fenómenos bajo leyes, atendiendo a sus regularidades. Esto admite una interpretación determinista (física) e indeterminista (naturalista). De cualquier modo, en ambos casos, una de las principales motivaciones es la búsqueda de un marco conceptual para explicar la agencia moralmente responsable.<sup>6</sup> Con vistas a ello, se establecen distintos tipos de relaciones entre la libertad y el orden causal: se pueden reconciliar (compatibilismo, determinismo blando) o aislar (determinismo duro, escepticismo, libertarismo).

### **Compatibilismo: del análisis condicional a las propuestas de Frankfurt y Strawson**

El compatibilismo es la posición que articula la libertad de la voluntad con el determinismo aunque, en su versión contemporánea, esto no implica que los compatibilistas presenten argumentos a favor del determinismo. Más bien la idea es que, de concederse que el determinismo es verdadero, no representa una amenaza para la libre voluntad y por lo tanto para la moral. Podemos ver los primeros antecedentes del compatibilismo en la filosofía de los estoicos, desde una cosmovisión que conciliaba el destino y las acciones humanas. En

---

6 Cabe señalar que existen distintas formas de comprender la naturaleza de la responsabilidad moral. Una de las principales distinciones se la debemos a Susan Wolf (1990) y a Gary Watson (1996). Ambos observan dos aspectos de la responsabilidad constitutivos de la vida moral de los individuos, aunque difieren en cuál es el que prevalece. Estos son: la responsabilidad como atribución (*attributability*) o como «rendición de cuentas» (*accountability*). El primero refiere a la responsabilidad que se deriva del sistema de valores del agente, en quien se reflejan sus vicios y virtudes (enfoque aretaico). El segundo, el que tiene más peso en el debate libertad—determinismo, refiere a las demandas entre unos y otros, y da lugar a la aprobación, la culpa y el merecimiento.

la época medieval, filósofos de la talla de Agustín y Aquino presentaron teorías compatibilistas. En la Modernidad, Hobbes y Hume han hecho grandes contribuciones teóricas afianzando esta posición. Pero en lo que concierne a la perspectiva metaética de la problemática, comienza con el precursor de esta misma disciplina: George E. Moore. Este filósofo retoma en la corriente analítica una idea que ya estaba presente en el compatibilismo clásico moderno, a saber, que lo contrario a la libertad de voluntad son las restricciones externas, y esta idea toma la forma de un análisis condicional o hipotético. Según este planteo, somos libres al realizar una acción *A* en lugar de *B*, en virtud del hecho de que podríamos haber hecho *B* si así lo hubiésemos elegido. De este modo la libertad no requeriría actuar de forma contraria a la motivación más fuerte. Si el pasado hubiera sido diferente (incluyendo deseos, intenciones y/o elecciones) el agente podría haber actuado diferente. La libertad es compatible con el determinismo porque no excluye la posibilidad de que la acción *A* haya sido determinada por una elección. Así, Moore adopta una postura conciliatoria entre determinismo y responsabilidad moral fundada en un compatibilismo analítico.

Algunas de las principales críticas a esta posición ponen en tela de juicio el carácter condicional mismo, sugiriendo un análisis categórico (por ej., Austin, 1954). Esta clase de objeciones apuntan al corazón del compatibilismo, que descansa en la idea de que el antecedente *si* expresa una condición hipotética. Para solucionar este problema, por ejemplo, Davidson (1973) propone otro tipo de lectura, considerando que el antecedente no refiere a una acción sino a un estado (creencia o deseo). Así, mantiene la idea compatibilista de que la libertad es un poder causal de los agentes.

A pesar de los esfuerzos, el análisis condicional ha sido puesto en tela de juicio en la misma tradición compatibilista a partir del texto de Harry Frankfurt *Alternate Possibilities and Moral Responsibility* (1969). La formulación condicional expresaba una idea intuitiva: la libertad de voluntad —y con ello la responsabilidad moral— requiere alternativas de acción. Es lo que Frankfurt llama «Principio de Alternativas Posibles» (PAP). Sin ir más lejos, la tesis de Frankfurt es que en ciertos contextos este principio es irrelevante para atribuir responsabilidad. Para sostener su idea distingue entre las causas que explican la acción (razones y motivos del agente) y los factores que la hacen inevitable, que no necesariamente la causan. Para ello presenta un estilo de ejemplos que se han consolidado como casos paradigmáticos (*Frankfurt-style examples*):<sup>7</sup>

---

<sup>7</sup> Frankfurt mismo sostiene que la estructura del siguiente ejemplo es lo suficientemente flexible para admitir diversos casos (cfr. 1969:834).

Supongamos que alguien —digamos Black— desea que Jones<sub>4</sub> realice una acción determinada. Black está dispuesto a ir lejos para salirse con la suya, pero prefiere evitar mostrar su mano innecesariamente. Entonces, espera a que Jones<sub>4</sub> esté a punto de decidir qué hacer, y no hace nada hasta que esté claro para él (Black es un excelente juez para tales cosas) que Jones<sub>4</sub> va a decidir hacer algo *distinto* de lo que él quiere que haga. Si se torna claro que Jones<sub>4</sub> va a decidir hacer otra cosa, Black toma medidas efectivas para hacer que Jones<sub>4</sub> decida hacer, y haga, lo que él quiere que haga. Cualesquiera sean las preferencias e inclinaciones iniciales de Jones<sub>4</sub>, luego, Black se saldrá con la suya. (Frankfurt, 1969:835, cursiva del autor).

Aquí Frankfurt muestra que Jones<sub>4</sub> dispone de un solo curso de acción y, si Black no interfiere, podemos afirmar que es responsable de su conducta. A pesar de la amenaza de la intervención, Jones<sub>4</sub> puede llevar a cabo sus acciones por motivos propios. De esta forma, a la vez que muestra la irrelevancia del PAP, Frankfurt pone en el centro al agente como *fuerza* de la acción, al ser esta causada por sus razones y decisiones.

Lo anterior se articula con otro de sus textos (Frankfurt, 1971), el cual también ha sido clave para la reflexión acerca de la agencia. Allí el autor distingue niveles de deseos, pues observa que mediante la reflexión, los deseos de primer orden (por ejemplo, querer hacer algo) pueden dar lugar a otros de segundo orden (*querer* querer hacer algo). Cuando estos últimos expresan lo que el agente efectivamente quiere hacer, decimos que reflejan su voluntad. Por ejemplo, un adicto puede querer consumir drogas y tener un deseo irrefrenable, y a su vez ser consciente del impacto negativo para su salud, de modo que quisiera no tener ese deseo. En cambio, él quisiera ser conducido por el deseo de dejar las drogas, a pesar de que siente que las necesita. De modo que se identifica con sus *voliciones*, es decir, sus deseos «más altos» (*higher-order*), no con los que lo guían *prima facie*. Para Frankfurt el agente es libre cuando logra articular o «engranar» (*mesh*) su voluntad con las voliciones que quiere que lo conduzcan, logrando cierta armonía. Así, las acciones compulsivas como la adicción (correspondientes al primer nivel), no expresan siempre los verdaderos deseos o creencias de los agentes; pueden ser, por el contrario, impedimentos internos para la autonomía o autodeterminación. Con este aparato teórico, Frankfurt permite concebir el control sobre las propias acciones de una forma que no era posible en el compatibilismo del análisis condicional.

La otra gran concepción contemporánea de la vertiente compatibilista, es la inaugurada por *Freedom and Resentment* (1962) de Peter F. Strawson. A diferencia de Frankfurt, Strawson no rechaza la concepción de libre voluntad basada en alternativas posibles, ya que observa que algunas ideas asociadas

a dicha visión reflejan un aspecto vital para la regulación de la conducta en formas socialmente deseables. Strawson plantea dos puntos de vista, no excluyentes entre sí. Por un lado, afirma que desde una perspectiva relativa a las relaciones interpersonales (que llama perspectiva «participante») es posible experimentar «actitudes reactivas» dirigidas hacia las personas y sus acciones. Estas pueden ser positivas: gratitud, amor, aprobación; o negativas: resentimiento, culpa, indignación, etc.; e indican que premiamos o condenamos acciones de terceros —o también propias—. Por otro lado, desde otro punto de vista (que denomina «objetivo») se toman en consideración las afectaciones psicológicas de los agentes. Esto expresa una distancia ante las reacciones mencionadas anteriormente, lo que permite observar las acciones desde un punto de vista desinteresado, puramente teórico. En consonancia, los agentes no son considerados plenamente responsables y son vistos como sujetos cuya conducta puede ser tratada, para, por ejemplo, curarla, evitarla o entrenarla. La distancia frente a las actitudes reactivas no implica una posición determinista; pero como compatibilista Strawson cree que si el determinismo es verdadero, no socava la libertad. Además, nuestras actitudes reactivas son naturales e irrevocables; más aún, son constitutivas de nuestra humanidad. A diferencia de algunos de sus antecesores (por. ej., Schlick, 1930) sostiene que nuestro sistema de premios y castigos debería adoptar una justificación «en retrospectiva» (*backward-looking*), que no ignore estas actitudes, aunque considera la atribución de responsabilidad como parte de prácticas comunitarias. Así, esta forma de compatibilismo se aleja de los esquemas metafísico-conceptuales (también del racionalismo de Frankfurt) y descansa en la psicología moral junto con las prácticas que la contemplan.

### **Incompatibilismo: libertarismo, determinismo duro y escepticismo**

Aunque, como mencionamos, los compatibilistas no siempre admiten como verdadero el determinismo, sus tesis son articulables con una de sus versiones, caracterizada como «determinismo blando», expresión hallada por ejemplo en Schlick (1939). Sin embargo, sus teorías excluyen otro tipo de determinismo: aquel que sostiene el «argumento de la consecuencia» o «argumento consecuencial» (*consequence argument*), una de las principales estrategias a favor del incompatibilismo. Según este razonamiento, si mantenemos una posición determinista, debemos aceptar que nadie podría tener incidencia en los hechos del futuro, puesto que tales eventos se hallan determinados por el pasado y por las leyes de la naturaleza, factores sobre los cuales nadie tiene

incidencia.<sup>8</sup> La idea es que si el argumento es verdadero, la única forma de concebir agentes libres es abrazando el incompatibilismo, más precisamente el libertarismo.

Los defensores del libertarismo, entonces, afirman que hay libre voluntad y excluyen el determinismo. Esta posición halla su antecedente más remoto en el epicureísmo. Según los libertarios, el individuo ejerce su agencia en cuanto se determina por sí mismo y no por causas externas. Si bien generalmente se reconocen tres tipos de libertarismo (agente causal, evento-causal y no causal), me centraré aquí en una de estas concepciones: la del agente causal. En cierto sentido puede comprenderse a Kant como un exponente de ella.<sup>9</sup> En la filosofía contemporánea, uno de los teóricos que respalda esta postura es Roderick Chisholm. Recordemos que desde el enfoque determinista un evento se considera causado solo por otro evento, constituyendo así una cadena causal. Frente a esto, Chisholm sostiene que no solo los eventos pueden constituir causas, sino también los agentes, cuyas acciones no pueden retrotraerse y reducirse a eventos, sino que son producto del ejercicio de su libre voluntad. Desde esta perspectiva, el agente es considerado sustancia, fuente de toda acción. Cada uno es como un «primer motor inmóvil» en el sentido de que es causa sin estar determinado por otras causas a su vez (cfr. Chisholm, 1964:32). Este enfoque permite una idea más robusta de la agencia humana que las visiones compatibilistas del análisis condicional, porque ser agente no consiste solo en poder haber hecho las cosas de otro modo, sino en la determinación de la propia voluntad, que da lugar a la posibilidad de haber *querido* (*willed*) de otro modo. Chisholm comparte con Frankfurt la relevancia del agente como fuente de la acción, sin embargo depende de un incompatibilismo metafísico que distingue a los agentes de otros tipos de entidades.

Otras formas de incompatibilismo son resultado de la negación de la libertad de la voluntad y de la afirmación del determinismo (duro) o incluso del indeterminismo. Sin embargo, a diferencia de los libertarios los siguientes incompatibilistas adoptan una posición escéptica respecto a la libertad de la voluntad y/o la moral.<sup>10</sup> El determinismo duro —o escéptico— halla en Spi-

---

8 Una de las expresiones más famosas de este principio es la de Van Inwagen: «Si el determinismo es verdadero, entonces nuestros actos son consecuencia de las leyes de la naturaleza y los eventos del pasado remoto. Pero lo que sucedió antes de que nacióramos no depende de nosotros, como tampoco depende de nosotros qué son las leyes de la naturaleza. Por lo tanto, la consecuencia de aquellas cosas (incluyendo nuestros actos presentes) no depende de nosotros» (1983:56).

9 Esto es, en tanto Kant concibió la libertad de los agentes de un modo que trasciende el orden determinista de la naturaleza. Chisholm adhiere a la lectura libertaria de Kant (cfr. 1964:33). Sin embargo hay que señalar que también se admite una interpretación compatibilista (Hudson, 1994).

10 A propósito de esto, es interesante lo que han puesto de manifiesto algunos filósofos contem-

noza uno de sus antecedentes clásicos más conocidos y aunque en la actualidad no cuenta con muchos adeptos, uno de los más reconocidos es Ted Honderich. Este filósofo propone una metaética no cognitivista y sostiene que los juicios referidos a la responsabilidad moral, como el resto de los juicios morales, no poseen valor de verdad. Coherente con los avances científicos actuales, concede que a un micronivel es posible sostener el indeterminismo (tal como postula la mecánica cuántica). Aun así, en la dimensión de la vida cotidiana, nos dice, solo es posible admitir el determinismo, regido por las leyes causales (incluyendo la actividad cerebral): «Es todo causa y efecto, en lo que podríamos llamar vida real» (Honderich, 2002:5).

Un importante apoyo a la posición escéptica —que no pertenece al canon filosófico pero es ineludible— proviene de los experimentos llevados a cabo por el neurocientífico Benjamin Libet en torno a la consciencia sobre nuestras decisiones. Este trabajo ofrece un enfoque precursor y ha marcado el camino de otras investigaciones sobre este tema (por ej., Soon *et al.*, 2008). El descubrimiento de Libet muestra la existencia de procesos cerebrales específicos que preceden temporalmente (en 350–400 ms) la conciencia de la propia intención de actuar. Libet llamó a dichos procesos «potencial de preparación» (*readiness potential*). Según sus conclusiones esto mostraría que las intenciones o decisiones conscientes no son las verdaderas causas de las acciones voluntarias, puesto que hay actividades neurofisiológicas inconscientes que son secuencialmente anteriores. No es difícil notar que las conclusiones de su experimento pueden interpretarse como una confirmación del determinismo y una amenaza a la libre voluntad (como la lectura que realiza Mele, 2013). Sin embargo, suele pasarse por alto que esto no llevó a Libet a afirmar la verdad del determinismo —tampoco del indeterminismo— porque, sostiene, no tenemos pruebas científicas suficientes para ello (Libet, 2002). En cambio, apuesta a la realidad de la libertad de voluntad, que se nos muestra como un fenómeno evidente y no como una mera ilusión (retomando lo dicho en la segunda sección, adopta una posición subjetivista y realista en este tópico —cfr. Libet, 2002:563—). De cualquier modo, los avances en neurociencias y otras ramas (psicología, genética, fisiología, etc.) nos obligan a repensar nuestras concepciones de la conciencia y del control sobre nuestras acciones, poniendo en crisis el carácter fuertemente racionalista de la agencia humana.

Otras teorías escépticas son las que se conocen como «sucesoras del determinismo duro», entre ellas las de Derk Pereboom (también podemos nombrar a Galen Strawson, 1994). Del determinismo duro —como el de Hon-

---

poráneos (como Strawson y Pereboom), esto es, que el escepticismo referido a la libertad de voluntad no conlleva necesariamente la adopción de la misma postura respecto a la moral en general.



derich— retoman el escepticismo frente a la libertad de voluntad, pero a diferencia de este, no sostienen la verdad del determinismo. Pereboom se proclama agnóstico ante dicha cuestión (cfr. Pereboom, 2014:3). Pero el núcleo de la teoría de este autor es el rechazo de la libre voluntad y de la responsabilidad moral basada en el merecimiento (*desert-based*) asociada a dicha noción:

Que un agente sea moralmente responsable de una acción en este sentido, es que sea suya de tal manera que merecería ser culpado si entendiera que fue moralmente incorrecto, y merecería ser elogiado si entendiera que fue moralmente ejemplar. El merecimiento en cuestión aquí es básico en el sentido de que el agente merecería ser culpado o elogiado sólo porque realizó la acción, en virtud de una comprensión de su estatus moral y no meramente, por ejemplo, a causa de consideraciones consecuencialistas o contractualistas. (Pereboom, 2014:2)

Este modo de entender la responsabilidad exige una concepción de voluntad en un sentido fuerte de control, que comprenda las acciones como dependientes del agente. En esta concepción, que es la que domina el debate, libertad de voluntad y responsabilidad moral (basada en el merecimiento) se sostienen o caen en simultáneo. Esto no evita que los teóricos escépticos ofrezcan otra concepción de la responsabilidad moral, «prospectiva» (*forward-looking*), por ej. consecuencialista o contractualista. Estos enfoques distan de los retrospectivos principalmente porque la reprobación de un acto no tiene como objetivo el castigo o la culpa basados en el resentimiento o la indignación, sino que apunta a la prevención de la acción en el futuro. Esto no significa negar el fenómeno de las actitudes reactivas ni que se abogue por la supresión de todas las emociones morales.

Otro elemento para destacar de este teórico es que ofrece uno de los argumentos recientes más importantes contra el compatibilismo (Pereboom, 2014 – cap. 4). Allí acusa a los compatibilistas de intentar desligar de responsabilidad al agente solo en situaciones donde este está visiblemente sujeto a factores externos, como los ejemplos de manipulación, pero lo que a su juicio no observan es que siempre es posible retrotraer las acciones a factores más allá de su control. Por último, es interesante la idea de que la negación del binomio libertad de voluntad–responsabilidad moral basada en el merecimiento no cancela nuestras concepciones de moral, agencia y tampoco de sentido de la vida.

## Los referentes en el debate actual

Finalmente, podemos mencionar algunas corrientes que han cobrado relevancia en las últimas décadas y cuyos antecedentes más inmediatos son las teorías que hemos repasado. Teniendo en cuenta el panorama abierto a partir de los años 60, no es de sorprender que las tendencias sean muy diversas. En lo que concierne a las teorías compatibilistas, podemos mencionar el *nuevo disposicionalismo*, que tiene como principal exponente a Kadri Vihvelin (2004, también Smith, 2003). Vihvelin vuelve a la clásica idea de reconciliar el determinismo causal con las alternativas de acción y de elección, pero evita el enfoque condicional, poniendo en el centro las habilidades de los agentes. Otra visión que se ha extendido entre los compatibilistas es la que refiere a la respuesta del agente ante las razones (*reason-responsive*). Aquí se concibe al individuo como sensible a las consideraciones racionales, afirmando que es libre cuando responde a ellas. La teoría más acabada es la de Fischer y Ravizza (1998) y entre las más recientes podemos mencionar la de Carolina Sartorio (2016). Por otra parte, siguiendo la línea de Strawson, uno de los abordajes más originales proviene de David Shoemaker (2017), quien propone una teoría basada en la respuesta del agente (*reason-dependent*) a partir de una analogía entre las emociones implicadas en la responsabilidad moral y las reacciones de humor.

En cuanto al incompatibilismo, se han continuado desarrollando las distintas versiones del libertarismo: del lado de la vertiente del agente causal, podemos mencionar a Randolph Clarke (2003), quien, en la misma línea de Chisholm, apuesta al carácter irreducible del agente como fuente de sus acciones. Por otra parte, para el libertarismo de eventos (aquel que sostiene que la causa de las acciones radica en los eventos) una de las referencias obligadas es Robert Kane (1996) y también podemos mencionar a Laura Ekstrom (2000). Por último, respecto a la variante no causal —la menos extendida—, el principal exponente es Carl Ginet (1990). La otra forma de incompatibilismo que vimos es el determinismo duro cuyo último exponente fue Honderich pero su espíritu escéptico reaparece en las concepciones sucesoras del determinismo duro. Pereboom es uno de los principales exponentes en la actualidad del denominado «escepticismo optimista», en contraposición podemos mencionar a Saul Smilansky (2000) quien sostiene que la libertad de voluntad es solo una ilusión (es en este sentido que ha sido calificado como pesimista) pero es necesaria en la práctica.

Por último, podemos mencionar que el debate continúa desarrollándose con las llamadas visiones «no estándar», generando nuevas intersecciones de las distintas categorías. Es así como, por ejemplo, podemos asistir a variantes como el *compatibilismo libertario*, que propone Christian List (2019), quien además ha puesto el foco en la agencia grupal.

## Referencias bibliográficas

- Aristóteles (2007). *Ética Nicomaquea* (trad. de Eduardo Sinott). Colihue.
- Austin, J.L. (1956). Ifs and Cans. *Proceedings of the British Academy*, 42, 109–132.
- Ayer, A.J. (1954). Freedom and Necessity. *Philosophical Essays*. Macmillan.
- Chisholm, R. (1982). Human Freedom and the Self. En Watson, G. (Ed.), *Free Will* (pp. 24–35). Oxford University Press.
- Clarke, R. (2003). *Libertarian Accounts of Free Will*. Oxford University Press.
- Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, 60, 685–700.
- Double, R. (2002). Metaethics, Metaphilosophy and Free Will Subjectivism. En Kane, R. (Ed.), *The Oxford Handbook of Free Will* (pp. 506–528). Oxford University Press.
- Ekstrom, L.W. (2000). *Free Will: A Philosophical Study*. Westview Press.
- Fischer, J.M. y Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Frankfurt, H. (1969). Alternate Possibilities and Moral Responsibility. *Journal of Philosophy*, 66, 829–839.
- Frankfurt, H. (1971). Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, 68, 5–20.
- Ginet, C. (1990). *On Action*. Cambridge University Press.
- Haji, I. (2009). *Incompatibilism's Allure: Principle Arguments for Incompatibilism*. Broadview Press.
- Hoefer, C. (2016). Causal determinism. En Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>
- Honderich, T. (2002). *How Free Are You?: The Determinism Problem*. Oxford University Press.
- Hudson, H. (1994). *Kant's Compatibilism*. Cornell University Press.
- Kane, R. (1996). *The Significance of Free Will*. Oxford University Press.
- Libet, B. (2002). Do We Have Free Will? En Kane, R. (Ed.), *The Oxford Handbook of Free Will* (pp. 551–564). Oxford University Press.
- List, Ch. (2019). *Why Free Will is Real*. Harvard University Press.
- McKenna, M. y Pereboom, D. (2016). *Free Will: A Contemporary Introduction*. Routledge.
- Mele, A. (2006). *Free Will and Luck*. Oxford University Press.
- Mele, A. (2013). Free Will and Neuroscience. *Philosophic Exchange*, 43, 1–17.
- Moore, G.E. (1912). *Ethics*. Oxford University Press.
- Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. Oxford University Press.
- Sartorio, C. (2016). *Causation and Free Will*. Oxford University Press.
- Schlick, M. (1939). When is a Man Responsible? En Schlick, M., *Problems of Ethics* (pp. 143–156). Prentice-Hall.
- Shoemaker, D. (2017). Response-dependent responsibility; or, a funny thing happened on the way to blame. *The Philosophical Review*, 126(4), 481–527.
- Smilansky, S. (2000). *Free Will and Illusion*. Oxford University Press.
- Smith, M. (2003). Rational Capacities. En Stroud S. y Tappolet C. (Eds.), *Weakness of Will and Varieties of Practical Irrationality* (pp. 17–38). Oxford University Press.
- Soon, C.; Brass, M.; Heinze, H. y Haynes, J. (2008). Unconscious Determinants of Free Decisions in the Human Brain, *Nature Neuroscience*, 11(5), 543–545.

- Strawson, G. (1994). The Impossibility of Ultimate Responsibility? *Philosophical Studies*, 75(1–2), 5–24.
- Strawson, P. (1962). Freedom and Resentment. *Proceedings of the British Academy*, 48, 1–25.
- Talbert, M. (2016). Moral Responsibility. En Zalta, E. (Ed.), *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/moral-responsibility/>.
- Van Inwagen, P. (1983). *An Essay on Free Will*. Clarendon Press.
- Vihvelin, K. (2004). Free Will Demystified: A Dispositional Account. *Philosophical Topics*, 32, 427–450.
- Watson, G. (2004). Two Faces of Responsibility (1996). En Watson, G., *Agency and answerability: Selected essays* (pp. 260–288). Oxford University Press.
- Wolf, S. (1990). *Freedom within Reason*. Oxford University Press.

## 15. El problema de la suerte moral

Gustavo A. Beade

El vínculo entre aquello que no controlamos y el modo en que juzgamos esas acciones constituye la llamada *suerte moral*, sencillamente ser alabado o reprobado en parte por la mera suerte. La suerte y el control, básicamente, son conceptos contrapuestos. De este modo, todo aquello que no está bajo mi control se encuentra librado a la suerte. Muchos filósofos y teóricos han realizado grandes esfuerzos para construir una noción del control para limitar la suerte. De este modo, ellos pretenden establecer qué es lo que está bajo mi control y qué es lo que depende de la suerte. Creo que es poco persuasivo el intento de construir la idea del control como un límite de la suerte. Así, también, entiendo que es bastante complejo poder establecer qué está bajo nuestro control y qué no lo está. Sin embargo, no voy a afirmar que no podemos controlar nada de lo que hacemos porque reduciría nuestra vida a una visión fatalista. Esta visión diría que no somos responsables por nada de lo que hacemos porque todo está librado a cursos causales externos.<sup>1</sup>

Mi objetivo es presentar en este breve trabajo algunos aspectos acerca de la suerte y también una introducción a la suerte moral. En primer lugar, describo algunas de las formas con las que aludimos a la suerte en nuestra vida comunitaria. Presento en esta descripción algunas intuiciones sobre esta cuestión y las muestro en contraposición a los planes individuales y colectivos así como también a los intentos de controlar nuestras acciones. Esto es, vale decir, a la forma y dirección que queremos darle a nuestra vida. Más adelante, en la segunda parte del texto, voy a relacionar estas intuiciones con la forma en la que juzgamos moralmente algunas de nuestras conductas.

### ¿Por qué la suerte?

Habitualmente, nos referimos a la suerte en diferentes sentidos. Puede ser que destaquemos nuestra buena o mala suerte. En algunas ocasiones, pensamos que se relaciona con aquello que ocurre inesperadamente y que, en definitiva, termina beneficiándonos en algún sentido: *¡qué suerte que encontré este billete tirado en el piso!*; en otras ocasiones hacemos referencia a acciones que redundan en alguna satisfacción o, incluso, un beneficio: *qué suerte que salió*

---

1 Esta descripción en Malamud Goti (2004).

*mi número en la lotería*; pero, por último, también hay hechos que decidimos no realizar y que con ello, evitamos un momento desagradable: *qué suerte que no fui a esa fiesta*. También nos referimos a la suerte en aquellas circunstancias en las que las cosas nos suceden: *por suerte* leí un libro relacionado con las preguntas y pude contestarlas bien en el examen de literatura; o *tuve la mala suerte* de enfermarme justo antes de ir a rendir el examen.

Sería posible reconocer que estas, como otras situaciones, son cosas que nos suceden y que tendemos a explicarlas de una forma vinculada a la suerte. Habría diferentes maneras de encontrar dinero y la suerte, como un factor determinante en sí mismo, perdería entidad si pudiéramos atribuir esta circunstancia a una acción o una omisión intencional de alguna persona en concreto. Por otra parte, a veces es difícil poder determinar si una consecuencia es atribuible o no a la suerte. Así, afirmar que si encuentro dinero en una vereda se debe a que, seguramente, alguien lo perdió, se le cayó o quizá lo arrojó intencionalmente para que alguien más necesitado que él cubriera sus carencias económicas. Esto podría solo ser, en algunos casos, una explicación posible. Si encuentro el dinero en un saco que hace tiempo no utilizo, fue mi propia omisión de revisar los bolsillos al quitármelo lo que origina que el dinero esté aún ahí. Es en circunstancias como esas donde la apreciación de la suerte podría ser alterada, pero en definitiva, no atinamos a negar su existencia para explicar todos esos hechos. De este modo, no sería solo la suerte de encontrar el dinero tirado en la vereda, sino de encontrar dinero antes de que otra persona lo hubiera encontrado; es la suerte de encontrar el dinero en el bolsillo de un saco viejo antes de que decidiera donar mis prendas usadas o viejas a una entidad de caridad o que el gobierno decidiera mediante una ley modificar la denominación de la moneda que acabo de encontrar.

Podríamos suponer que transitar por una calle vacía a las tres de la mañana o, incluso, caminar mirando hacia abajo pensando en el destino del mundo, nos darían más chances de encontrar cosas perdidas. Esto supondría, en algún caso, una explicación plausible para el hecho de encontrar algún objeto de valor tirado en una vereda. Sin embargo, usualmente habría dos opciones para explicar esta situación: «¡encontré dinero tirado en la calle!» o «¡qué suerte, encontré dinero tirado en la calle!». Probablemente habría una forma algo compleja e incierta de explicar que «...iba caminando a las tres de la mañana rumbo al trabajo como todos los días, doblé en la avenida y, mientras iba pensando si seguiría o no trabajando en este lugar, encontré un billete tirado en la vereda». En este punto, mi explicación se podría completar diciendo que fue mi propia acción de caminar a esa hora hacia el trabajo, mirando hacia abajo, lo que derivó en el hallazgo del billete. Entiendo que, apelando a la mera causalidad, sería posible explicar muchos de estos hechos, aunque no

sin mayores problemas. Así, por un lado, no es posible, a partir de allí, dar razones para justificar el hecho en concreto. En verdad, la ausencia de esas razones es lo que en parte podría mostrar que el hecho está influido por la suerte.

Por otra parte, no es así como explicamos este tipo de circunstancias. Notablemente, preferimos reducir la cuestión a un acto azaroso, inesperado, a un hecho producto de la *suerte*. Al parecer, no hay nada que haya o no haya realizado que pudiera modificar ese momento en que solo estaba ahí para que el hallazgo del dinero ocurriese «por obra y gracia» de la suerte. Si hubiera hallado el billete en una avenida muy transitada habría demasiadas variantes por considerar y sería imposible explicarlo de una forma tan detallada. Así, estar allí en ese momento, advertir justo en ese instante que tenía el cordón desatado, arrodillarme a atarlo y encontrar un billete que *pudo haberlo* encontrado alguna de las cientos de personas que transitaban por ahí en ese mismo momento. En definitiva, la explicación causal que elija para relacionar con esta acción compleja, será tan indeterminada y a la vez, tan arbitraria que no tendría más certezas que si solo lo concibiera como un hecho producto de la suerte. Si, finalmente, gano el premio mayor de la lotería, muy probablemente también lo atribuya a la suerte y no a otro tipo de explicación más compleja. Así, más allá de que hice algo (aposté a un número en la lotería) no puedo explicar el hecho de que hoy gané el premio y no lo hice la semana pasada. Alguien, sin embargo, podría intentar explicar que si apuesto una cierta cantidad de dinero para ganar otra mayor, en definitiva, ganar o no ganar el premio depende de una cuestión de probabilidades. Estas aumentarían a medida que disminuye la cantidad de participantes o que transcurra un tiempo en el que siga apostando a la misma combinación de números. Sin embargo, lo cierto es que habría muchas cosas que considerar que están más allá de la propia acción de elegir algunos números y apostar dinero. También allí se busca en la suerte algo que no es explicable, sencillamente, por la aplicación de la ley de probabilidad. Creo que la probabilidad no puede explicar por qué salió mi número y no otro; por qué salió hoy y no ayer, o el lunes que viene; la probabilidad tampoco puede explicar el hecho de quienes ganan la lotería en su primera apuesta o el de quienes apuestan su vida entera sin ganar nunca el premio. Seguramente, habría expertos que podrían, a través de intrincados procesos, explicar de alguna forma plausible estos procesos aunque siempre luego de ocurrido el resultado.

No es uno de mis objetivos aquí discutir cuáles son las formas de explicación que apelan a la causalidad ni tampoco a la posibilidad de utilizar la probabilidad para determinar comportamientos futuros. Asumo aquí la apelación a la suerte como una forma de explicar las cosas que hacemos, y las cosas que nos suceden, dado que es así como en comunidades como en las que vi-

vimos se explican determinadas circunstancias. Este punto de vista al que llamo *optimista*, tiene la particularidad de creer en la influencia de la suerte en estas y otras cosas que hacemos. Para volver al ejemplo anterior, no es posible explicar el hecho de ganar la lotería por mi única intervención. Existe una innumerable cantidad de factores que contribuyen al resultado, que no pueden ser identificados sino a través de la simplificación que nos otorga recurrir a la suerte. En definitiva, necesitamos recurrir a la suerte para explicar el hecho concreto: haber ganado la lotería. Es por eso que quien fuera el afortunado ganador diría: «*¿Qué suerte que salió mi número!*».

En definitiva, incluso los aspectos más triviales en los que invocamos a la suerte, están relacionados con tres aspectos básicos de nuestra vida: la responsabilidad, la distinción entre lo que hacemos y lo que ocurre más allá de nosotros, y la finalidad con la que actuamos, esto es, los planes que rigen nuestras vidas. Nuestros planes, deseos e intenciones se concentran en las acciones que llevamos a cabo. Estas acciones se ven, en muchos casos, alteradas por causa de factores que están más allá de nuestros cálculos que, además, no podemos explicar claramente. Ante la ausencia de esta explicación, atribuimos estos hechos sencillamente a nuestra buena o mala suerte. Asumir estas circunstancias implica sostener lo que llamo una versión *ingenua* de la suerte.

Quisiera ahora centrarme en otros aspectos menos triviales respecto de la suerte, que se encuentran más allá de estas simples intuiciones comunitarias. Me interesa señalar que hay muchos más factores que muestran la presencia de la suerte en nuestra vida que los que podemos fácilmente apreciar. De este modo, parece claro que encontrar dinero en la vereda y ganar la lotería dependen de la suerte. Mi punto es admitir la intromisión de la suerte en muchas de las cosas que hacemos. Esta circunstancia se presenta como algo más difícil de reconocer, dado que implica admitir que tenemos menos control sobre las cosas que hacemos o sobre los planes que construimos.

## **El control, las planificaciones y la influencia de la suerte**

La suerte es bastante más amplia en su influencia de lo que, en principio, estaríamos dispuestos a advertir o conceder. Existe una determinada cantidad de factores cotidianos que están altamente influidos por la suerte. En general, tratamos de organizar nuestra existencia a partir de limitar el azar en cuestiones sencillas. Dejar la ropa preparada para el día siguiente, comprar la comida para tener durante todo el mes, tener dinero ahorrado para evitar pasar por situaciones difíciles son intentos para limitar la influencia de la suerte. Veamos esto un poco más en detalle.



Desde que somos niños y niñas, nos enseñan que es necesario ir a la escuela porque aprender y estudiar es importante para nuestro futuro. Estos argumentos continúan en algunos casos al terminar la escuela secundaria y se orientan a construir la idea de que el estudio disminuye la posibilidad de transformarnos en desempleados. La idea sería: «hay que estudiar para tener mayores posibilidades en la vida» según la hemos escuchado de padres, madres y parientes. Estudiar constituiría, en este aspecto, la balsa sobre la que nos sostendríamos en caso de sufrir un hecho inesperado, en definitiva, de tener que superar un momento de mala *suerte*.

Es notable como esta imagen podría ser trasladada hacia otros aspectos de nuestra vida cotidiana. Los seguros de vida constituyen otro ejemplo de lo que se supone que intentamos controlar: la posibilidad de morir inesperadamente. Admitimos, entonces, que no podemos controlar nuestra vida frente a un deceso no predecible. Aceptar esta posibilidad nos genera una cierta angustia que trasciende nuestra propia persona y se traslada hacia nuestros seres queridos. En verdad, nos angustia la posibilidad de que nuestra familia quede desamparada ante una muerte inesperada. La intención detrás de obtener un seguro de vida es retomar un poco ese control: el saber que nada está a la deriva, incluso luego de mi desaparición física. En definitiva, muchos de nuestros proyectos están limitados y terminan descarrilados por la suerte. Sin embargo, somos reticentes a admitir que muchas de las actividades que llevamos adelante con la finalidad de neutralizar la suerte son infructuosas. De hecho, lo más difícil es convencernos de que nuestra vida está invadida completamente por la suerte y que controlamos menos cosas de las que creemos. Así, nuestros juicios evaluativos acerca de las consecuencias de nuestras conductas podrían ser equivocadas.

Todo el tiempo estamos condicionados por factores externos que, al parecer, están bajo nuestro control y que, en realidad, no lo están: salí temprano para encontrarme con una amiga pero, de imprevisto, antes de salir hacia la cita, una discusión con un compañero de estudios me afectó de tal manera que lo único que quiero es irme a mi casa y estar solo. El hecho de que me haya dejado llevar por mis pasiones, que mi ira hacia mi colega haya sido más fuerte que mi compromiso anterior con mi amiga, es solo una muestra de que el control sobre las situaciones cede antes muchas circunstancias inesperadas. Las emociones, los deseos, determinadas reacciones como la indignación y el resentimiento dependen de circunstancias que en muchos casos, escapan a nuestro control. De este modo, estas circunstancias inesperadas también influyen sobre nuestros juicios morales. Así mi amiga que quedó esperando en el bar, difícilmente apruebe moralmente mi comportamiento. Hay muchas cosas que quizá ella debería considerar, como mi fuerte temperamento o mi

conocida tendencia a resolver las cosas impulsivamente. Pese a ello, mi amiga lo único que considerará es que se quedó esperando, que yo no llegué a la cita y que, en consecuencia, no sería tan *buen* amigo. Mi amiga tendrá dudas de volver a querer encontrarse conmigo.

Quisiera ampliar en lo que sigue la influencia de la suerte en nuestros juicios morales de responsabilidad y empezar a introducir ciertas nociones vinculadas con la suerte moral.

## **Nuestros juicios morales y la suerte**

Una gran cantidad de las cosas que hacemos dependen de factores ajenos a nuestra voluntad y también de nuestras complejas planificaciones. Existen circunstancias en nuestra vida en donde la suerte juega un papel determinante. Esto ocurre, por ejemplo, con relación a nuestros juicios morales. La influencia de la suerte en todo lo que hacemos modifica nuestros juicios morales de diferentes maneras. En comunidades como en la que vivimos, evidentemente los resultados de las acciones que llevamos a cabo son muy relevantes. La aparición de la suerte, en muchos casos, altera la manera en que enjuiciamos ciertos cambios que ocurren en el mundo. En general, la influencia de la suerte es imperceptible para nosotros que decidimos concentrarnos en las consecuencias, en lo que en definitiva modificó el mundo. Sin embargo, también en nuestras comunidades acostumbramos a reconocernos a partir de esas modificaciones del mundo, por las que, en mayor o menor medida somos responsables.

La *suerte moral* se presenta en nuestras comunidades de diferentes formas, pero particularmente en la forma en la que distribuimos premios y castigos. La vinculación de la suerte con las consecuencias de lo que hacemos. Algunos creen que es posible evitar la suerte concentrándonos solo en nuestras acciones y no en sus resultados. Según algunos filósofos y teóricos nuestras acciones son libres de la suerte: son las consecuencias lo que no podemos controlar. Este aspecto nos lleva a la difícil situación de separar lo que está influido por la suerte y lo que no lo está. Esta discusión se basa en determinar si existen acciones que caen bajo nuestro control y otras que están por fuera de él. Hay ciertas cosas que hacemos en las que la suerte no juega ningún papel apreciable: compré el diario por la mañana porque quiero enterarme cómo sigue el mundo en el que vivo; acepté el trabajo porque necesito dinero para poder pagar el alquiler de casa; no iré a tu fiesta porque no deseo encontrarme con mis ex-compañeros. Todas estas son acciones u omisiones sencillas que han sido decididas por nosotros y por las que debemos responder sin transgredir

reglas de etiqueta o normas morales: debo pagar el diario al vendedor para poder llevármelo, debo cumplir con ciertos requisitos de forma y de fondo para poder mantener el trabajo y responder ante mi amigo por ausentarme de su festejo. La suerte no impone de por sí una excusa o una justificación para todas nuestras conductas que luego resultan indeseadas o desagradables. En definitiva uno podría evaluar en sus acciones la influencia de la suerte, pero no podría esgrimir la suerte como un excluyente de responsabilidad a partir del resultado ocurrido. Lo que ocurre luego de que decidimos tomar un curso de acción no siempre es predecible. Sin embargo, no es posible separar las consecuencias del curso de acción que elegí.

Ahora bien, esta responsabilidad que debemos adjudicarnos por nuestras conductas, implica también la adjudicación, en muchos casos, de un premio o un castigo por ese hecho (Rivera López, 2002:76). En este sentido, la mera suerte, que puede influir en nuestras conductas habituales, se transforma en *suerte moral*, es decir, en el hecho de que una persona sea moralmente censurada o elogiada, en parte, por su suerte.

## La suerte moral

Mostré que mucho de lo que hacemos depende de factores ajenos a nuestra voluntad y que también escapan a nuestras complejas planificaciones. Entre otros factores, la suerte juega también un papel determinante en nuestros juicios morales. Esta influencia en lo que decidimos hacer es lo que llamamos suerte moral.

La idea de la suerte moral, según Bernard Williams, muestra a la moral en sí misma como un valor supremo, en el sentido kantiano, que no puede ser alterada por factores contingentes. Así, nuestros juicios morales sobre mérito, responsabilidad e inculpación no deben estar condicionados por la suerte (Williams, 1993:251). En este punto, quien quisiera defender una posición kantiana, en la cual la moral ocupe un lugar supremo, debería dejar de lado a la suerte. Esto es así porque una moral kantiana *es* inmune a la suerte (Williams, 1981:21). La moral valora a la voluntad con independencia de cómo repercute en el mundo la acción correspondiente. De esta manera, la suerte no tiene ningún lugar en la evaluación de aquello que decidimos hacer. Es el propio Kant quien estableció que la moral está regida por la buena voluntad con independencia de la manera en que la acción ocurra como el agente quiera. Además, según Kant, la buena voluntad es buena en sí misma y por lo tanto se impone ante circunstancias vinculadas con la naturaleza o el destino (Kant, 1785:393-394). Sin embargo, sostiene Williams, que esta noción kan-

tiana de la moral podría decepcionarnos, dado que es difícil negar la influencia de la suerte en nuestros juicios morales.<sup>2</sup> Este es, precisamente, el punto que me interesaría desarrollar aquí.

En los primeros trabajos dedicados a la suerte moral, Bernard Williams (1981) y Thomas Nagel (1979) escribieron serias consideraciones poniendo en duda la fuerte inferencia de la ética kantiana de que la moral no puede ser permeable a la suerte. Para seguir esta línea trazada por Williams y Nagel, creo que la suerte modifica nuestros juicios morales en modos diferentes. En lo que sigue, voy a presentar tres de las «clases» de suerte que fueron desarrolladas por Thomas Nagel en su trabajo *Moral Luck*.<sup>3</sup> Así, me voy a referir a la *suerte constitutiva*, a la *suerte situacional* y a la *suerte en el resultado*.

Es importante destacar que estos tres conceptos no pueden entenderse parcialmente. Existe una vinculación estrecha entre los tres tipos de suerte que enuncié. En efecto, no puede concebirse la *suerte en el resultado*, esto es, las consecuencias de mis acciones, sin considerar previamente de dónde provengo (*suerte constitutiva*), por qué estoy en *este* momento y lugar, llevando a cabo *esta* acción (*suerte situacional*). Estas circunstancias personales crean las condiciones para esta acción. En este trabajo, sin embargo, me voy a concentrar en los problemas que se presentan en la suerte en el resultado. La discusión sobre la suerte constitutiva y la suerte situacional es abordada desde hace un tiempo por algunos teóricos interesados en la filosofía política. Veamos a continuación los esfuerzos llevados a cabo por estos filósofos para intentar neutralizar la suerte.

## La suerte constitutiva y la suerte situacional

Nuestro origen depende de nuestra propia *suerte constitutiva*. Es así que la suerte determina donde nacemos, los padres que tenemos, y en parte, también las inclinaciones que nos mueven en la vida. De este modo, desde escribir con la mano derecha o la izquierda, y tener afición por los deportes violentos es algo que no elegimos y con lo que convivimos desde un comienzo.<sup>4</sup>

---

2 La cuestión aquí es establecer si la moral es un orden normativo supremo. Si esto es así, la suerte no tiene un lugar posible. Según sostiene Williams esta interpretación de la moral no es correcta.

3 Dejo de lado la cuarta categoría de suerte utilizada por Nagel, que tiene que ver con los condicionamientos que tenemos para actuar sobre la base de las circunstancias precedentes. Me enfocaré solamente en tres de los tipos de suerte que se enumeran en su trabajo.

4 Estas tempranas diferencias que los seres humanos debemos atravesar ha llevado a autores kantianos como John Rawls, a aislar la suerte originaria de los individuos a partir de la «posición original» y detrás de lo que llama el «velo de ignorancia». Allí, desde una posición en donde quienes están allí, desconocen sus gustos, sus intereses, sus capacidades, es posible que puedan decidir sobre cuestiones morales desde el punto de vista más imparcial posible. Ver Rawls (1971:118 y ss.).

Implica el tipo de persona que somos, esto es, no solo una cuestión acerca de lo que intencionadamente hacemos sino también acerca de nuestras inclinaciones, capacidades y temperamento (Nagel, 1979:28). En este sentido, no podemos saber si seremos sensibles, afectivos, altos, bajos, etc. (Malamud Goti, 2004:861).

Por otra parte, quienes somos, nuestro carácter y nuestros gustos, nos ubican en ciertos lugares y momentos que en parte, responden a la llamada *suerte situacional*. Esta modalidad de la suerte tiene vinculación con el tipo de problemas y situaciones que uno debe enfrentar (Nagel, 1979:28). Básicamente, las situaciones que rodean nuestra vida y el medio en el cual actuamos. Algunas personas nacieron dentro de una familia y una comunidad armoniosa y ordenada, otros, en cambio, se han criado en pobres y violentos vecindarios (Malamud Goti, 2004:863).

Desde hace algunos años, hay un sector de la filosofía política que ha entendido que es necesario tomar medidas para neutralizar, de alguna forma, la (mala) suerte constitutiva que provoca serias diferencias, difíciles de recomponer, entre los distintos agentes. Estas deficiencias se trasladan y forman parte de nuestra suerte situacional. Se trata de los filósofos comprometidos con el llamado liberalismo igualitario. Este pensamiento tiene, entre sus miembros más notables, a John Rawls (1971), Ronald Dworkin (2000), G.A. Cohen (1989), y el mismo Thomas Nagel (1991), entre otros. Ellos son quienes han sido denominados en un trabajo, ya clásico, como «igualitarios de la suerte» (Anderson, 1999). Se ha dicho que los autores citados tienen la visión de que el punto fundamental del igualitarismo reside en compensar a las personas por su inmerecida mala suerte: el haber nacido con una dotación pobre, malos padres, una personalidad desagradable, que hayan sufrido accidentes o enfermedades, entre otras cosas. Sin embargo, el trabajo llevado a cabo por estos «igualitarios de la suerte» no es uniforme. En este sentido, existen muchas diferencias entre sus propuestas que buscan modificar estos efectos del azar.

Pese a los esfuerzos realizados, estos «igualitarios de la suerte» no han podido acordar aún muchas cuestiones. Para dar tan solo un ejemplo, no existe un acuerdo acerca de si deben neutralizarse solo las diferencias económicas, o si también es necesario hacerlo respecto de los talentos naturales. En su caso, cómo llevar a cabo esta parificación también es problemático. Una opción que se ha presentado, y que trato de explicar aquí a grandes rasgos, es reducir, en alguna medida, las riquezas de las personas más acaudaladas de la comunidad para beneficiar a las personas más desaventajadas. Esto debería ser así, dicen estos filósofos, porque la suerte los benefició y por eso no hay méritos detrás de su riqueza que, en definitiva, es por eso moralmente cuestionable.

En estos debates no se ha logrado acordar si, garantizando un piso mínimo de paridad económica, debemos darnos por satisfechos o si, por el contrario, también es necesario neutralizar el talento natural de una persona. Estas desigualdades que luego —según el caso— podrían transformarse en una diferencia económica.<sup>5</sup> En efecto, hay un círculo de discusión sobre estos aspectos de la suerte que son más bien amplios lo cual indicaría que la suerte moral, al menos, en dos aspectos: suerte constitutiva y suerte situacional, son considerados —en algún sentido— por la filosofía política. Esta preocupación tiene una correlación en la discusión de la suerte en el resultado que, sin embargo, transita por otros carriles. Por el momento me gustaría decir algo más sobre esta idea y los problemas que se genera admitir la intervención de la suerte en las cosas que hacemos.

### La suerte en el resultado

Una cuestión adicional y quizá la más discutida entre filósofos y juristas es la influencia de la suerte en las consecuencias de los actos que llevamos a cabo, la llamada *suerte en el resultado*. La suerte en el resultado radica en cómo se concretan nuestros proyectos, como culminan nuestras acciones en el mundo, luego de que intencionadamente movemos el cuerpo (Malamud Goti, 2004:863). Es la buena o mala suerte que tenemos en las consecuencias de aquello que hacemos luego de que decidimos tomar un curso de acción (Statman, 1997). Como señalé antes, la suerte es aquello que es ajeno a nuestra voluntad y que se nos presenta en la vida desde el comienzo. No es posible decidir dónde nacer, si seremos sanos, ni tampoco qué talentos tendremos (o si tendremos alguno). La suerte excede el hecho puntual, el infortunio o la *mala suerte*. Gráficamente, la suerte es como el agua, todo lo penetra (Katz, 2000:792). La reticencia en aceptar esta realidad puede verse en nuestras comunidades en las que los individuos intentan todo el tiempo modificar cursos azarosos en el afán de pretender controlar sus vidas. Cuando estos intentos fallan decidimos «echarle la culpa» al hecho en concreto, al cálculo imprevisto, a lo irregular o poco común. Intentamos construir la idea del infortunio como algo excepcional. Como señalé al comienzo, creo que esta for-

---

<sup>5</sup> Es claro que esta cuestión tampoco puede formularse sin alguna aclaración al respecto. En la discusión de neutralizar un talento natural tiene suma relevancia identificar qué talentos querríamos neutralizar. En apariencia, no habría motivos suficientes para intentar neutralizar el talento del futuro mejor jugador de squash del mundo si este deporte no fuera una actividad valorada en nuestra comunidad. Si este fuera el caso, esta desigualdad sería, a todas luces, irrelevante para los demás ciudadanos.

ma de entender aquello que nos importa y construir nuestro propio mundo es, al menos, engañosa. La suerte es más amplia, y debemos convivir con ella mal que nos pese.

Aceptar en alguna medida la idea de que la suerte interviene constantemente, que hay un factor de azar en nuestras decisiones, implica aceptar que —en gran parte— no todas nuestras conductas están controladas por la voluntad. También significa que nosotros no somos responsables de todo lo que hacemos porque aparecen situaciones en las que la suerte influye y limita seriamente nuestra capacidad de decidir qué cursos deberíamos tomar. En este sentido, se ha afirmado que «Algo que ocurre por causa de la suerte es algo que ocurre fuera del control de cualquiera» (Zimmerman, 1987:376). La cuestión aquí no es menor, estamos hablando de la posibilidad de poder controlar nuestras acciones o no. Algunos filósofos creen que se trata de una seria discusión acerca del grado de autonomía que rige nuestra vida (Zimmerman, 1987:375).

Teóricos y filósofos morales se ocuparon de discutir la suerte en el resultado desde distintos puntos de vista. Uno de los aspectos más desarrollados se vincula con las distinciones que efectuamos para la justificación moral de reproches y condenas morales y castigos criminales.<sup>6</sup> Habitualmente, filósofos morales y teóricos penales sostienen que es inapropiado censurar, condenar y castigar a un agente por eventos y estados de cosas que escapan a su dominio racional. Estos teóricos piensan que un agente sólo puede ser responsable si estuvo en condiciones de deliberar racionalmente: de pesar razones en favor de algún curso de acción. Esta consideración responde a la idea de que la acción responsable no es el resultado de impulsos incontrolados y, así, dominados por emociones que nos acosan sino que expresa una libre decisión del agente. Esta decisión debería ser considerada, incluso, si consistiera en dejarse llevar por propia y libre decisión por sus emociones o deseos (Malamud Goti, 2008:32). Para esta tradición, la responsabilidad se agota en la intención inicial de actuar y en la de perseverar en el curso elegido sin considerar lo que suceda en el mundo externo una vez que se desarrollan los hechos. Vuelvo a señalar que, para esta postura, nada de lo que ocurre más allá del control de la voluntad del agente debe pesar en la evaluación moral de cualquier acto. Así, acciones realizadas con los mismos fines y en idénticas circunstancias merecen igual responsabilidad aunque unas causen daños y las otras no (Malamud Goti, 2008:32). Denomino a esta posición la *tesis de la equivalencia*.<sup>7</sup>

---

6 Si bien hay una gran cantidad de filósofos del derecho penal que piensan que la responsabilidad moral y la responsabilidad penal son análogas, hay otros autores que afirman que esto se basa en un error. Ver Shoemaker (2013).

7 Así por ejemplo entre muchos otros esta posición es defendida por Alexander y Ferzan (2009), Ferrante (2009), Ashworth (1993), Sancinett (1991), Kaufmann (1954), Zielinski (1973).

Por otra parte, contra la *tesis de la equivalencia*, hay quienes alegan que es justo, razonable o equitativo adjudicar mayor responsabilidad a alguien que sí causa daños o pone en peligro a otros. Según esta posición, si un agente causa un daño o genera situaciones de peligro debe ser castigado más severamente que otro agente que realice la misma conducta pero que no concretó algún daño o peligro. A esto se lo ha denominado en varias oportunidades como la *tesis de la diferencia*.<sup>8</sup> Esta tesis lo que defiende es la idea de que en el mundo en el que vivimos, los resultados de las cosas que hacemos cuentan. Si esto es así, tanto en los premios o reconocimientos como en los reproches morales, no hay razones para pensar que es posible construir un sistema de responsabilidad penal que implique resolver esta cuestión de un modo diferente.<sup>9</sup>

### **Palabras finales**

En esta breve presentación he intentado resumir algunos puntos importantes para empezar a pensar el problema de la suerte moral. En el trabajo, me enfoqué en presentar algunas ideas sobre el modo en el que pensamos la suerte y la suerte moral. Por último exhibí algunas discusiones que se presentan en la responsabilidad moral, pero también en la responsabilidad penal en particular en lo referido a la llamada suerte en el resultado.

---

8 Esta posición es defendida entre muchos otros por Tadros (2011), Jakobs (2010), Malamud Goti (2008), Duff (1996), Stratenwerth (1975).

9 Sin embargo, ver Beade (2017).



## Referencias bibliográficas

- Alexander, L. y Ferzan, K. (2009). *Crime and Culpability, A Theory of Criminal Law*. Cambridge University Press.
- Anderson, E. (1999). What is the Point of Equality? *Ethics*, 109, 287–337.
- Ashworth, A. (1993). Taking the consequences. En Shute, S.; Gardner, J. y Horder, J. (Eds.), *Action and Value in Criminal Law* (pp. 107–124). Clarendon Press.
- Beade, G.A. (2017). *Suerte moral, castigo y comunidad: un análisis de la relevancia moral de la suerte en el resultado*. Marcial Pons.
- Cohen, G.A. (1989). On the Currency of Egalitarian Justice. *Ethics*, 99, 906–944.
- Duff, R.A. (1996). Subjectivism objectivism and the role of theory. En Simester, A. y Smith, A. (Eds.), *Harm and Culpability* (pp. 19–44). Clarendon Press.
- Dworkin, R. (2000). *Sovereign Virtue. The Theory and Practice of Equality*. Harvard University Press.
- Ferrante, M. (2009). Recasting the Problem of Resultant Luck. *Legal Theory*, 15, 267–300.
- Jakobs, G. (2010). Erfolgsunwert und Rationalität. En Joecks, W. et al. (Eds.), *Rechts–Wirtschaft–Strafe, Festschrift für Erich Samson zum 70. Geburtstag* (pp. 43–54), CF Müller.
- Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten*. Akademie Ausgabe IV.
- Katz, L. (2000). Why the Successful Assassin is More Wicked than the Unsuccessful One. *California Law Review*, 88, 791–812.
- Kaufmann, A. (1954). *Lebendiges und Totes in Bindings Normentheorie*. Verlag Otto Schwartz.
- Malamud Goti, J. (2004). Rethinking Punishment and Luck. *Tulsa Law Review*, 39, 861–874.
- Malamud Goti, J. (2008). *Suerte, moralidad y responsabilidad penal*. Hammurabi.
- Nagel, T. (1979). Moral Luck. En Nagel, T., *Mortal Questions* (pp. 24–38). Cambridge University Press.
- Nagel, T. (1991). *Equality and Partiality*. Oxford University Press.
- Rawls, J. (1971). *A Theory of Justice*. Belknap Press.
- Rivera López, E. (2002). Responsabilidad y suerte moral. Circunstancias y consecuencias de la acción. En Navarro, P. y Redondo, C. (Comps.), *La relevancia del derecho. Ensayos de filosofía jurídica, moral y política* (pp. 275–290). Gedisa.
- Sancinetti, M. (1991). *Teoría del delito y disvalor de acción. Una investigación sobre las consecuencias prácticas de un concepto personal de ilícito circunscripto al disvalor de acción*. Hammurabi.
- Statman, D. (1997). The Time to Punish and the Problem of Moral Luck. *Journal of Applied Philosophy*, 14(2), 129–135.
- Stratenwerth, G. (1975). Zur Relevanz des Erfolgsunwertes im Strafrecht. En Grünwald, G. et al. (Eds.), *Festschrift für Friedrich Schaffstein* (pp. 177–193). Verlag Otto Schwarz & Co.
- Shoemaker, D. (2013). On Criminal and Moral Responsibility. En Timmons, M. (Ed.), *Oxford Studies in Normative Ethics* 3 (pp. 154–178). Oxford University Press.
- Tadros, V. (2011). Obligations and Outcomes. En Cruft, R.; Kramer, M. y Reiff, M. (Eds.), *Crime, Punishment and Responsibility, The Jurisprudence of Antony Duff* (pp. 173–192). Oxford University Press.
- Williams, B. (1981). *Moral Luck*. Cambridge University Press.
- Williams, B. (1993). Postscript. En Statman, D. (Ed.), *Moral Luck* (pp. 251–258). State University of New York Press.
- Williams, B. (1995). Internal Reason and the Obscurity of Blame. En Williams, B., *Making Sense of Humanity* (pp. 35–45). Cambridge University Press.
- Zielinski, D. (1973). *Handlungs- und Erfolgsunwert in Unrechtsbegriff*. Duncker & Humblot.
- Zimmerman, M. (1987). Luck and Moral Responsibility. *Ethics*, 97, 374–386.

## 16. Neuroética, moral y el paso del «ser» al «deber ser»

*Daniel González Lagier*

El nacimiento de la neurociencia moderna suele situarse en 1888, con el descubrimiento de las neuronas por Santiago Ramón y Cajal. Posteriormente, Charles Sherrington analizó las conexiones entre ellas y Edgard Adrian registró actividad bioeléctrica en todo el sistema nervioso. Los avances en psicofarmacología a mediados del siglo xx y, sobre todo, la aparición de las técnicas de neuroimagen en los noventa y principios del siglo xxi, han dado un impulso decisivo a esta ciencia. Las técnicas de neuroimagen han permitido el estudio de los mecanismos que tienen lugar en el cerebro durante los procesos de adquisición del conocimiento y de la formación de estados mentales, dando lugar a la neurociencia cognitiva. Del tronco central de la neurociencia se han ramificado nuevas disciplinas, que pueden verse como aplicaciones de la misma a distintos ámbitos. En 2002 se celebró en San Francisco el congreso *Neuroética: esbozando un mapa del terreno* que consagraba esta nueva disciplina, dedicada a la aplicación de la neurociencia a los temas tradicionales de la ética, y en 2007 la MacArthur Foundation creó el proyecto «Derecho y Neurociencia» para reunir a varios conocidos neurocientíficos, filósofos y juristas de diversos países con el fin de profundizar en la intersección entre las neurociencias y el Derecho. Surgía así una importante línea de investigación que se enfrentaba, ni más ni menos, a la tarea de estudiar científicamente el fenómeno de la normatividad, esto es, la capacidad humana de darse normas, de actuar conforme a ellas y de valorar la conducta de los demás y adscribir responsabilidad a partir de ellas. En este breve trabajo trataré de dar cuenta de algunos de esos intentos y de las dificultades que, en mi opinión, deben afrontar.

### 1.

Hay dos maneras de entender la «neuroética» (Roskies, 2002; Bonete Perales, 2010:64; Cortina, 2010:131): (1) como *ética de la neurociencia*, una parte de la ética que trataría de establecer un marco ético para las investigaciones neurocientíficas y sus aplicaciones; y (2) como *neurociencia de la ética*, el estudio de la conducta ética desde el punto de vista de las investigaciones sobre el funcionamiento del cerebro en los procesos cognitivos y de razonamiento. Algunos problemas de la ética de la neurociencia son los siguientes: si está justificado o no el uso de los descubrimientos neurocientíficos para la mejora de las

capacidades mentales o sensoriales de los humanos (el llamado «transhumanismo»), en qué condiciones es legítimo el uso en los tribunales de pruebas basadas en técnicas neurocientíficas (como la prueba P300 o *brainfingerprinting*, que se espera que permita determinar si el sujeto miente observando las variaciones en las ondas cerebrales ante ciertos estímulos), qué valor en relación con la atribución de responsabilidad hay que conceder a determinadas disfunciones cerebrales, o si es correcto —y en qué casos— usar técnicas de control de la conducta basadas en conocimientos neurocientíficos. Y algunos problemas de la neurociencia de la ética son los siguientes: la discusión general sobre el libre albedrío y su relación con la responsabilidad, la cuestión de la fundamentación de las normas morales en emociones como la repugnancia, el análisis del papel de la oxitocina o de las llamadas «neuronas espejo» en nuestra conducta ética, o si la neurociencia puede fundamentar conclusiones normativas acerca de la corrección de nuestros juicios morales. En este trabajo la expresión «neuroética» se usa en el segundo sentido, esto es, como neurociencia de la ética.

## 2.

La neuroética intenta dar cuenta de la normatividad a partir de las ciencias empíricas, en la línea del «darwinismo moral» de Spencer, que trataba de fundamentar las normas éticas en el proceso evolutivo, o la «sociobiología» de E.O. Wilson, que pretendía explicar el comportamiento de todos los animales extrapolando el egoísmo genético del mundo de los genes a los demás ámbitos de la vida. En la actualidad, desde la neuroética muchos autores (J. Haidt, M. Hauser, P. Churchland, M.S. Gazzaniga,...) proponen dar cuenta de las opiniones y creencias morales como un conjunto de intuiciones, emociones y capacidades en gran parte innato, inscrito en nuestro cerebro por las «fuerzas» de la evolución:

Hemos desarrollado un instinto moral, una capacidad que surge en cada niño, diseñada para generar juicios inmediatos sobre lo que está moralmente bien o mal, sobre la base de una gramática inconsciente de la acción. Una parte de esta maquinaria fue diseñada por la mano ciega de la selección darwiniana millones de años antes de que apareciera nuestra especie; otras partes se añadieron o perfeccionaron a lo largo de nuestra historia evolutiva y son exclusivas de los humanos y de nuestra psicología moral. (Hauser, 2008:17)

Autores como Kant, Hume (o cierta interpretación del mismo) y George Moore han contribuido a que el mundo de las normas se haya concebido como un ámbito separado e independiente del de la naturaleza, pero la neurociencia ha aportado argumentos con los que se ha puesto en cuestión esta imagen de lo normativo, tratando de mostrar que la moral no tiene tal autonomía. La principal vía para esta «reducción de la normatividad» tiene relación con ciertas conclusiones a las que se ha llegado a partir de la realización de «experimentos éticos» de tipo psicológico y neurocientífico. Jonathan Haidt, por ejemplo, ha realizado una gran cantidad de encuestas en las que planteaba ciertos problemas morales, llegando a la conclusión de que las respuestas se basaban en rápidas intuiciones que luego los sujetos no sabían cómo racionalizar. A partir de ahí propuso que la moral es más una cuestión de intuiciones y emociones que de razones (Haidt, 2012). Posteriormente Joshua Greene examinó con técnicas de neuroimagen la actividad cerebral de sujetos a los que se le planteaban diversos dilemas morales, llegando a la conclusión de que en los problemas morales que afectaban más personalmente a los sujetos, había una mayor actividad de la zona que regula las emociones (Greene, 2012). Marc Hauser, también por medio de encuestas que apuntan a rasgos morales universales, ha propuesto la hipótesis de la existencia de un órgano moral, esto es, de una capacidad innata para el desarrollo de códigos morales que determina en parte el contenido de los mismos (algo así como una gramática moral universal, en analogía con la gramática universal postulada por Chomsky) (Hauser, 2008). En esta línea, William Casebeer ha afirmado que la teoría moral aristotélica de la virtud es más plausible desde un punto de vista neurobiológico que la teoría moral kantiana o la de John Stuart Mill.

El punto de partida de Casebeer es que cada una de estas teorías contiene implícita una psicología moral específica que exige capacidades cognitivas diferentes. Así, la teoría de Kant parecería requerir «al menos la capacidad de comprobar la consistencia lógica de máximas universalizadas de una manera independiente de la contaminación del afecto y la emoción», capacidad que se correspondería con las funciones de la región frontal del cerebro. La teoría utilitarista de Mill requiere la capacidad de realizar cálculos utilitarios y el cultivo de emociones que nos muevan a procurar la felicidad de los demás, lo que implicaría las regiones prefrontal, límbica y sensorial del cerebro. La ética aristotélica de la virtud, por último, sería la más exigente, porque requiere educar nuestro carácter de manera que nuestros apetitos se coordinen con las buenas razones; esto implica una «psicología global» que solicita una intervención coordinada de las regiones del cerebro anteriormente mencionadas. Pues bien, nuestro autor cree que hay pruebas para aceptar, tentativamente, que la «cognición moral» pone en marcha de manera coordinada diferentes

sistemas y redes cerebrales relacionadas tanto con la cognición como con las emociones (es decir, las regiones prefrontal, frontal, límbica y sensorial: lo que podría llamarse «la zona de la cognición moral»), lo que muestra que «existe una clara convergencia entre la neuroética contemporánea y la psicología moral aristotélica» (Casebeer, 2003).

Como vemos, de lo que se trata es de derivar conclusiones acerca de la moral a partir de descripciones empíricas, de explicar la moral a partir del funcionamiento del cerebro. Sin embargo, algunas propuestas no se quedan en un nivel meramente descriptivo y explicativo, sino que tratan de derivar conclusiones normativas (como parece ser el caso de Casebeer). En sus versiones más radicales, se trata de reducir los valores o las normas morales a regularidades naturales, a los valores o las pautas de comportamiento que surgen de nuestra naturaleza física, asumiendo su carácter normativo y construyendo de esta forma un *realismo moral naturalizado*. De acuerdo con esto, la moral puede descubrirse por medios empíricos y las normas no forman parte de un reino distinto de la naturaleza.

### 3.

Podemos caracterizar este «realismo moral neurocientífico» (que sostendrían autores como M. Hauser, 2008; P. Churchland, 2012; M. Gazzaniga, 2015...) a partir de los siguientes rasgos:<sup>1</sup>

- A) *Intuicionismo*: nuestro juicio moral está dominado, o fuertemente influido, o condicionado en gran medida por intuiciones rápidas, inconscientes, no intencionales, ajenas a la reflexión. Ante los dilemas morales, nuestra respuesta suele ser intuitiva, irreflexiva, al menos en un primer momento. Por tanto, la deliberación y discusión racional sobre los problemas morales no tiene relevancia genuina (o la tiene muy menor, o solo en los casos en los que no nos vemos afectados personalmente). Su papel es —en la mayoría de las ocasiones— o bien una racionalización *ex post*, o bien solo pretende causar en el oponente la misma intuición.
- B) *Emotivismo*: estas intuiciones morales dependen fundamentalmente de nuestras emociones o son la expresión de estas emociones, como demuestran las pruebas neurofisiológicas basadas en la observación de qué zonas del cerebro tienen mayor actividad en el momento en que nos enfrentamos a dilemas morales. La tesis de que las decisiones morales provienen de la emoción —manifestada como intuición— puede verse como un caso

---

1 Para un análisis más detenido véase González Lagier (2017).

especial de la tesis general que enfatiza el papel de las emociones en la toma de decisiones (de cualquier tipo). Damasio sugiere que el modelo clásico de toma de decisiones como el resultado de una ponderación de razones a favor o en contra de una u otra respuesta no es descriptivamente adecuado. Los seres humanos no deciden en realidad de esa manera (si así fuera, las decisiones racionales, simplemente, no serían posibles, porque no es posible procesar toda la información necesaria en el tiempo en el que disponemos para actuar). Las emociones nos ayudan a tomar decisiones racionales porque «filtran» las alternativas posibles en función de si en el pasado alternativas semejantes fueron «etiquetadas» como adecuadas (produjeron placer o satisfacción) o inadecuadas (produjeron frustración o dolor). Las emociones son, por tanto, «marcadores somáticos» de las posibles alternativas (Damasio, 2004:201 y ss.).

- C) *Innatismo*: se trata de intuiciones y emociones innatas, transmitidas genéticamente, aunque luego pueden ser moldeadas culturalmente: nacemos con valores, creencias o principios morales independientes en su origen del aprendizaje, de manera que la moral no es enteramente un producto cultural.
- D) *Evolucionismo*: las intuiciones morales se ven como mecanismos que la evolución ha seleccionado porque aseguran la supervivencia de la especie. Uno de los principios que estaría detrás de estas intuiciones, según muchos neuroéticos, sería el altruismo, la cooperación o el principio de beneficencia. Pues bien, en opinión de autores como Michel Ruse, el «altruismo biológico» (la necesidad de cooperación) es tan esencial para los seres humanos que «la naturaleza nos ha llenado de ideas sobre la necesidad de cooperar»:

Pensamos que debemos ayudar, que tenemos obligaciones para con los demás, porque tener estas ideas va en nuestro interés biológico. Pero desde una perspectiva evolutiva estas ideas existen sencillamente porque aquellos de nuestros antepasados que las tuvieron sobrevivieron y se reprodujeron mejor que los que no. En otras palabras, el altruismo es una adaptación humana, igual que lo son nuestras manos y ojos y dientes y brazos y pies. Somos morales porque nuestros genes, modelados por la selección natural, nos llenan de ideas sobre la conveniencia de serlo. (Ruse, 2004:672)

- E) *Normativismo*: ¿es posible construir una ética basada en el funcionamiento de nuestro cerebro? Esto es lo que parecen pensar algunos filósofos y neuroéticos. Patricia Churchland, por ejemplo, critica la idea de que la ciencia no puede decirnos cómo debemos vivir y afirma que al igual que la salud «es un ámbito en el que la ciencia puede enseñarnos, y ya lo ha hecho, gran parte de lo que deberíamos hacer», también en «el ámbito de la conducta social (...)

podemos aprender mucho de la observación común y de la ciencia acerca de las condiciones que favorecen la armonía y la estabilidad social, así como la calidad de vida individual» (Churchland, 2012). También Gazzaniga propone la construcción de una ética universal basada en el funcionamiento del cerebro (Gazzaniga, 2015). La idea, tal como la expresa (críticamente) Adela Cortina sería que «entre el mundo del ser natural y el del deber ser (los códigos morales) existiría un lazo adaptativo que prescribiría establecer como normas éticas aquellas conductas capaces de favorecer la supervivencia» (2010:137).

#### 4.

Cada una de las anteriores afirmaciones, sin embargo, debe verse como una hipótesis, pero no como una tesis bien establecida, ya que a cada una de ellas se le puede oponer importantes objeciones aún no respondidas.

La primera tesis, de acuerdo con la cual los juicios morales dependen más de la intuición que de la razón, parece establecer una oposición intuición/razón demasiado radical, esto es, parece estar presuponiendo que una excluye a la otra. Sin embargo, como Bunge ha puesto de manifiesto (Bunge, 2013 – cap. III.1.), la idea de «intuición» engloba muchos fenómenos distintos (modos de percepción, formas de imaginación, inferencias rápidas, a saltos o incompletas, capacidad de síntesis, capacidad de evaluación de una situación y de elección de las mejores alternativas,...) y varias de esas formas de intuición no pueden verse como opuestas o excluyentes de la razón: la inferencia rápida o incompleta es un razonamiento embrionario o primitivo y la «aprehensión sinóptica —como señala Bunge— no es un sustituto del análisis, sino un premio al análisis esmerado» (Bunge, 2013 – cap. III.1). Pero, además, los diferentes tipos de intuición —incluso aquellos que no pueden verse en sentido estricto como razonamientos, ni siquiera incompletos— vienen favorecidos especialmente por el ejercicio continuado del razonamiento, del análisis sobre un problema, de la experiencia en una actividad o de la dedicación al estudio de una disciplina. También podría verse la intuición, como ha propuesto Peter Gärdenfors (2005), como un tipo de conocimiento implícito y especialmente difícil de explicitar, por no estar estructurado lingüísticamente. En definitiva, de acuerdo con este argumento la contraposición entre intuición y razón sería falsa; la intuición solo se contrapone al razonamiento explícito o proposicional. Las intuiciones a las que se refiere la neuroética podrían ser de este tipo y la dificultad de dar razones que las justifiquen podría derivarse de la dificultad de acceder a ese conocimiento implícito y no proposicional, pero no de la ausencia de tales razones en la mente del sujeto.

Por lo que respecta a la relación entre el juicio moral y las emociones, también se presentan algunas dificultades. El argumento que está detrás de esta vinculación es el siguiente: por una parte, las imágenes cerebrales demuestran que cuando razonamos moralmente se activan intensamente zonas del cerebro relacionadas con la emoción; por otra, los test psicológicos demuestran que en la mayoría de casos resolvemos los dilemas morales de manera intuitiva. Ambas cosas deben estar relacionadas: las intuiciones, por tanto, surgen de las emociones, de la actividad de las zonas afectivas del cerebro. Pero esta manera de reconstruir las decisiones morales plantea, en mi opinión, varios problemas:

- a) En primer lugar, la evidencia que tenemos del papel de las emociones en la toma de decisiones es, en realidad, indirecta. Lo que los neurocientíficos pueden comprobar es que cuando nos enfrentamos a un dilema moral las zonas del cerebro relacionadas con las emociones se activan de una manera especialmente acusada. Pero esto no permite inferir que las emociones generan el juicio moral. Podría ser que fuera el juicio moral el que genera la emoción: por ejemplo, darse cuenta, tras un análisis, de lo injusto de una situación puede generar indignación; o, en un dilema moral, ser consciente de que cualquier solución causará un daño puede provocar pesar. De acuerdo con Aristóteles las emociones se relacionan con las creencias en una doble dirección: por un lado, muchas emociones son generadas por creencias; por otro, muchas emociones generan o modifican nuestras creencias (González Lagier, 2009:26 y ss.). La neuroética parece estar viendo solo esta segunda conexión. Además, parece haber evidencia de que las emociones se relacionan más con la motivación de la acción moral que con el juicio moral (parece que los individuos psicópatas, a pesar de sus déficits emocionales, tienen intuiciones morales semejantes a los individuos «normales», pero no se ven motivados por ellas).
- b) En segundo lugar, la noción de emoción como «marcadores somáticos» las convierte en «impulsos ciegos», meras sensaciones que carecen de contenido proposicional, por lo que solo es posible dar una explicación causal de su relación con los juicios morales (y no una explicación teleológica o basada en razones). Esta manera de entender las emociones y su relación con la moral se aleja de las concepciones de la emoción más extendidas hoy en día entre los filósofos y muchos psicólogos. Para estas concepciones (que, en buena medida, reivindican la concepción de las emociones de Aristóteles), debe distinguirse en ellas al menos tres dimensiones distintas: 1) Una dimensión cognitiva (en un sentido amplio, que incluye desde una creencia hasta una mera percepción); 2) una dimensión afectiva o puramente fenomenológica (la sensación de placer o dolor); y 3) una dimensión mo-



tivacional (una tendencia a la acción). Las teorías cognitivas de la emoción centran su atención en el primer aspecto, las teorías «somáticas» y mecanicistas de la emoción se centran en el segundo y las teorías conductistas en el tercer elemento. Por último, las teorías no reductivistas tratan de dar cuenta de todos los aspectos de las emociones. Si se identifican las emociones con el segundo aspecto quedan muchos problemas por resolver: a) no se logra explicar la posibilidad de emociones inconscientes o sin sensación; b) no se logra explicar que las emociones puedan formar parte de explicaciones racionales (teleológicas), y no solo causales, de la conducta; c) no se logra explicar la relación entre las emociones y las creencias; d) no se logra dar cuenta de la posibilidad de evaluar a las emociones como razonables o no (en función de la creencia subyacente); y e) sobre todo, incurrir —de nuevo— en un distanciamiento tajante entre las intuiciones de las que se dice que proceden nuestros juicios morales y la razón (González Lagier, 2009 – cap. II).

Tampoco está libre de objeciones la tesis —estrechamente vinculada a las anteriores— del innatismo de las opiniones morales. Los argumentos a favor del innatismo descansan, de nuevo, en el carácter automático de la respuesta y en la incapacidad de dar razones (Haidt), junto con la coincidencia de respuestas a pesar de la heterogeneidad de los encuestados (Hauser). Sin embargo, para asegurar que nuestras creencias morales son innatas habría que descartar completamente que el automatismo se deba a la aceptación de principios profundamente arraigados, pero transmitidos culturalmente. Como señala Adela Cortina comentando los experimentos de Haidt:

El hecho de que las personas encuestadas respondan de forma intuitiva, es decir, inmediata, automática, sin tener conciencia de cómo han llegado a formular el juicio, y que en muchas ocasiones no sepan dar razón de por qué una acción les parezca buena o mala puede muy bien explicarse porque lo han aprendido socialmente y no lo han sometido a revisión. (2011:86)

Muchas normas de la moral social (como el incesto, en el ejemplo de Haidt, si es que se acepta su carácter moral) son heredadas del entorno social, que las inculca la mayoría de las veces sin dar razones que las justifiquen. Por otra parte, la universalidad de las normas o principios también debe tomarse con precaución: muchos de los supuestos comportamientos morales universales no son, en realidad, de carácter moral (o exclusivamente de carácter moral): por ejemplo, como hemos visto, Hauser ha postulado la universalidad del principio de cuidado de las crías (que, obviamente, tiene una clara expli-

cación evolutiva), pero ¿deberíamos extraer también la conclusión de que la búsqueda de alimento o el huir de los animales predadores tiene también valor moral? (Barta, 2013 – cap. III) ¿O la reproducción? Lo que quiero decir es que muchos hábitos aportados como ejemplos de conductas universalmente aceptadas no tienen (o no tienen exclusivamente) carácter moral.

Hemos visto que otras de las tesis características de este intento de dar cuenta de la moral desde un punto de vista biológico y neurocientífico es el darwinismo moral. Las intuiciones morales se ven como mecanismos que la evolución ha seleccionado porque aseguran la supervivencia de la especie. Esta tesis, sin embargo, adolece de cierta ambigüedad. Para aclararla conviene distinguir, de nuevo, entre la pretensión (descriptiva) de explicar la moral y la pretensión (de carácter normativo) de justificarla. A la primera podemos llamarla darwinismo moral descriptivo y a la segunda darwinismo moral prescriptivo. El darwinismo moral descriptivo, a su vez, puede tratar de explicar *la capacidad* del ser humano para tener una conducta ética (para evaluar las conductas como correctas o incorrectas desde el punto de vista moral) o tratar de explicar (lo que es más ambicioso) *el contenido* de la moral, esto es, tratar de explicar por qué creemos que ciertas conductas determinadas son correctas o por qué algunos principios o valores están tan extendidos (Ayala, 2013:61). El darwinismo moral descriptivo recurre a la idea de que comportarse moralmente o adecuar el comportamiento a ciertos principios es un rasgo que ha facilitado la evolución de la especie y su supervivencia. Las tesis normativas añaden que, puesto que eso es así, tales principios están justificados.

Respecto de las tesis explicativas, creo que debe afirmarse que se trata de hipótesis no suficientemente establecidas. Respecto de la tesis normativa, creo que es directamente el resultado de varios errores.

Tomemos la afirmación según la cual lo que explica la capacidad humana de evaluar conductas como buenas o malas y de ajustar la conducta a determinados principios es que esta capacidad es una ventaja evolutiva. Para aceptarla concluyentemente esta tesis debe rechazar la hipótesis alternativa (igualmente plausible) planteada por Francisco Ayala según la cual el comportamiento ético no es directamente un resultado de la evolución, sino solo indirectamente y en la medida en que es una consecuencia del desarrollo de la inteligencia humana; es decir, lo que tiene valor adaptativo y ha sido favorecido por la evolución es la inteligencia humana, no el hecho de ser capaz de comportarse moralmente (que es consecuencia a su vez de la inteligencia humana) (Ayala, 2013:66). Si, por el contrario, lo que se afirma es que los códigos morales vienen determinados por la evolución, el problema es que no parece encontrarse un conjunto de principios relevantes, que no estén formulados de una manera excesivamente vaga y vacía, que sean realmente uni-

versales; y, además, es posible encontrar tipos de conducta, como la agresividad o la territorialidad, que son importantes evolutivamente y no pueden ser aceptados como ejemplos de conducta moral. La moral que se deriva de la evolución podría ser terrible.

Por su parte, la tesis que trata de reducir lo que debemos hacer a aquello que es bueno para la supervivencia de la especie se enfrenta a la obvia objeción de que viola la ley de Hume: pasa de descripciones acerca de lo que es evolutivamente útil a prescripciones acerca de qué debemos hacer o cómo debemos vivir.

Como es sabido, en un famoso pasaje, Hume afirma:

En todo sistema moral del que haya tenido noticia, hasta ahora, he podido siempre observar que el autor sigue durante un cierto tiempo el modo de hablar ordinario, estableciendo la existencia de Dios o realizando observaciones sobre los quehaceres humanos, y, de pronto, me encuentro con la sorpresa de que, en vez de las cópulas habituales de las proposiciones: es y no es, no veo ninguna proposición que no esté conectada con un debe o no debe. Este cambio es imperceptible, pero resulta, sin embargo, de la mayor importancia. En efecto, en cuanto que este debe o no debe expresa alguna nueva relación o afirmación, es necesario que ésta sea observada y explicada y que al mismo tiempo se dé razón de algo que parece absolutamente inconcebible, a saber: cómo es posible que esta nueva relación se deduzca de otras completamente diferentes. (1981:689–690)

Este pasaje se ha interpretado habitualmente (aunque no unánimemente) como la prohibición de inferir de enunciados «sobre lo que es el caso» enunciados «sobre lo que debe ser». Por ello, los autores que defienden esta postura tratan de argumentar contra la validez de la ley de Hume. Examinemos con más detalle estos argumentos.

## 5.

Me parece que los argumentos que usualmente se esgrimen contra la ley de Hume suelen ser de dos tipos:

1) El primer tipo es el de los argumentos basados en contraejemplos: una manera frecuente de mostrar que es posible fundar normas o valores en descripciones consiste en presentar ejemplos de argumentos en los que aparentemente se realiza esta derivación. Esta es la estrategia seguida, en un famoso artículo, por John Searle (1980:178–201). Entre los neuroéticos también se ha recurrido a este tipo de argumentos. Marc Hauser, por ejemplo, propone el siguiente:

*Hecho*: la única diferencia entre un médico que aplica anestesia a una criatura y otro que no se lo aplica es que, sin ella, el niño sufrirá enormemente en el curso de una operación. La anestesia no tendrá ningún efecto pernicioso para el niño, sino que le hará perder temporalmente la conciencia y la sensación de dolor. Luego despertará, una vez acabada la operación, sin ninguna secuela negativa y con mejor salud, gracias al trabajo del médico.

*Juicio valorativo*: por consiguiente, el médico debe administrar anestesia al niño (2008:28)

Sin embargo, este tipo de argumentos parecen incurrir en uno de los siguientes errores: o bien confunden lo que es bueno o debido desde un punto de vista técnico con lo que es bueno o debido desde un punto de vista moral o normativo, o bien presentan como un argumento completo lo que en realidad es un argumento entimemático que incluye una premisa oculta, que es precisamente la norma o la valoración de la que se deriva la conclusión.

Para evitar el primer error, hay que advertir que no siempre que un enunciado incluye el término «deber» es un genuino enunciado normativo. A veces, «debe» expresa una conjetura («debe ser así» puede significar «probablemente es así»), otras veces, se puede sustituir por «tiene que» y expresa una necesidad práctica. Es importante distinguir entre deberes deónticos o genuinos y deberes técnicos, prudenciales o necesidades prácticas. Muchos de los ejemplos que se ofrecen como derivación de «debe» a partir de «es» no concluyen genuinos deberes deónticos, sino necesidades prácticas. Como señala von Wright, podemos:

encontrar dos respuestas principales a la cuestión de por qué una cierta cosa debe o puede o no tiene que ser hecha. Una es que existe una norma ordenando o permitiendo o prohibiendo la realización de esa cosa. La otra es decir que los fines y las conexiones necesarias hacen (o no) la realización u omisión de esa cosa una necesidad práctica. (1963:74)

Por lo que respecta al segundo error, es fácil darse cuenta de que muchas veces los argumentos que formulamos en contextos cotidianos no incluyen todas sus premisas. Es incluso factible pensar que en algunos casos es imposible en la práctica enunciar todas las premisas necesarias para llegar a la conclusión. Pero de ello lo que se sigue es la derrotabilidad o revisabilidad de la conclusión, no que su corrección no dependa de la premisa implícita. El ejemplo de Hauser presupone, para su corrección, una premisa según la cual se debe evitar el sufrimiento innecesario, que es una premisa normativa.

2) El segundo tipo de argumentos restringe la Ley de Hume a los argumentos deductivos: se afirma que lo que la Ley de Hume proscribía es derivar deductivamente un deber ser a partir del ser, pero que existen otro tipo de inferencias aceptables, como la inducción o la inferencia a la mejor explicación, por medio de las cuales sí se puede pasar de descripciones de hechos a normas (Casebeer, 2003:482; Churchland, 2012 – cap. 1). Una manera de defender que la Ley de Hume se refiere exclusivamente a inferencias deductivas consiste en entenderla como una consecuencia del principio de conservación de la lógica: en una deducción no es posible que se concluya algo que no estuviera incluido ya en las premisas. Las deducciones pueden hacer que seamos conscientes de un dato nuevo, pero este se encontraba ya en las premisas. Por ello, solo de proposiciones descriptivas no podemos deducir enunciados de deber ser. Esto ocurre con cualquier cosa. Como observa Pigden «existe un salto similar entre las conclusiones sobre los erizos y las premisas que no hacen mención de ellos. No se pueden obtener conclusiones sobre “erizos” a partir de premisas carentes de erizos (al menos no solo por la lógica)» (Pigden, 2004:570). Ahora bien, a diferencia de las deducciones, las inducciones y las abducciones sí amplían nuestro conocimiento, por lo que no rige para ellas el principio de conservación. Por tanto, si la Ley de Hume es solo una manifestación del principio de conservación de los argumentos deductivos, entonces tienen razón los que sostienen que esta no es aplicable a las inferencias no deductivas.

Pero, ¿es solo eso la Ley de Hume? Probablemente, no. Puede sostenerse que entre enunciados descriptivos y enunciados normativos existen importantes diferencias —a veces se habla de un «abismo lógico» entre ellas, o entre hechos, por un lado, y normas y valores, por otro—: así, los enunciados descriptivos tienen una *dirección de ajuste*<sup>2</sup> descendente (esto es, palabras a mundo: se pretende que las palabras se ajusten al mundo) mientras que los enunciados normativos tienen una dirección de ajuste ascendente (esto es, mundo a palabras: se pretende que el mundo se ajuste a las palabras). Los enunciados descriptivos son verdaderos o falsos, mientras que las normas o valores no lo son. Los enunciados que expresan deberes presuponen un punto de vista interno, en el que los términos deónticos (obligatorio, prohibido) son usados, mientras que si una descripción se refiere a una norma o a un deber, lo hace desde un punto de vista externo en el que los términos deónticos son solo mencionados. Todas estas diferencias entre descripciones y normas hacen que las primeras no puedan servir de razones, ni expresar razones, para justificar las segundas. No es solo que una justificación deductiva requiera que entre las

---

2 Una explicación de la idea de «dirección de ajuste» puede encontrarse en Searle (1976).

premisas esté aquello que se quiere deducir, es que —aun cuando se admita que la inducción o la abducción puedan tener alcance justificatorio— ningún enunciado descriptivo puede, por sí solo, ser una razón que justifique un enunciado prescriptivo. Puede aportar una razón explicativa de por qué aceptamos ciertas normas o valores; puede ser también una razón explicativa o, incluso, justificatoria, de otros enunciados descriptivos. Pero si se quiere concluir la justificación de una norma a partir de descripciones y por medio de argumentos no deductivos, se tiene la carga de la prueba. Y es interesante observar que los neuroéticos no lo han hecho.

## 6.

Si las anteriores consideraciones son correctas, cuando se propone que debemos seguir aquellas pautas de conducta que tienen un valor adaptativo, o bien simplemente estas se recomiendan como medidas prudentiales para mantener la supervivencia de la especie humana, pero entonces no tienen carácter moral, o bien se asume que la supervivencia de la especie es un fin moralmente valioso, en cuyo caso la normatividad no viene de los hechos, sino de esta asunción valorativa. ¿Quiere decir todo lo anterior que la neurociencia no puede aportar nada relevante para la comprensión de la normatividad? Esta sería una conclusión equivocada. Todo razonamiento genuinamente deóntico (esto es, si no se trata de necesidades prácticas o deberes técnicos) tiene una premisa normativa y una premisa fáctica. La neurociencia puede contribuir al establecimiento de esta premisa fáctica; y, más aún, si asumimos el principio «debe implica puede», puede establecer límites a las normas que tiene sentido establecer: esto es, la neuroética podría ayudarnos a entender cuál es el espacio de la moralidad. Pero no puede justificar por sí sola respuestas a los problemas éticos.

## Referencias bibliográficas

- Ayala, F. (2013). *Evolución, ética y religión*. Universidad de Deusto.
- Barta, R. (2013). *Cerebro y libertad. Ensayo sobre la moral, el juego y el determinismo*. Fondo de Cultura Económica.
- Bonete Perales, E. (2010). *Neuroética práctica. Una ética desde el cerebro*. Desclée de Brouwer.
- Bunge, M. (2013). *Intuición y razón*. Debolsillo.
- Casebeer, W.D. (2003). Moral cognition and its neural constituents. *Nature Reviews/Neuroscience*, 4, 840–846.
- Churchland, P. (2012). *El cerebro moral. Lo que la neurociencia nos cuenta sobre la moralidad*. Paidós.
- Cortina, A. (2010). Neuroética: ¿Las bases cerebrales de una ética universal con relevancia política? *Isegoría, Revista de Filosofía Moral y Política*, (42), 129–148.
- Cortina, A. (2011). *Neuroética y neuropolítica. Sugerencias para la educación moral*. Tecnos.
- Damasio, A.R. (2004). *El error de Descartes*. Crítica.
- Gärdenfors, P. (2005). Las intuiciones como conocimiento incompleto. *Revista de Filosofía*, XLIII(108), 51–55.
- Gazzaniga, M.S. (2015). *El cerebro ético*. Paidós.
- González Lagier, D. (2009). *Emociones, responsabilidad y Derecho*. Marcial Pons.
- González Lagier, D. (2017). *A la sombra de Hume. Un balance crítico del intento de la neuroética de fundamentar la moral*. Marcial Pons.
- Greene, J.D. (2012). Del «es» neuronal al «debe» moral: ¿cuáles son las implicaciones morales de la psicología moral neurocientífica? En Cortina, A. (Ed.), *Guía Comares de Neurofilosofía práctica* (pp. 149–158). Comares.
- Haidt, J. (2012). El perro emocional y su cola racional: un enfoque intuicionista social del juicio moral. En Cortina, A. (Ed.), *Guía Comares de Neurofilosofía práctica* (pp. 159–215). Comares.
- Hauser, M.D. (2008). *La mente moral. Cómo la naturaleza ha desarrollado nuestro sentido del bien y del mal*. Paidós.
- Pigden, Ch. R. (2004). El naturalismo. En Singer, P., *Compendio de ética* (pp. 567–580). Alianza Editorial.
- Roskies, A. (2002). Neuroethics for the New Millenium. *Neuron*, 35(1), 21–23.
- Ruse, M. (2004). La significación de la evolución. En Singer, P., *Compendio de ética* (pp. 667–680). Alianza Editorial.
- Searle, J. (1976). Una taxonomía de los actos ilocucionarios. *Teorema*, I(1), 43–77.
- Searle, J. (1980). *Actos de habla*. Cátedra.
- Von Wright, G.H. (1963). *The Varieties of Goodness*. Routledge & Kegan Paul.

## 17. Motivación moral. El debate entre el internalismo y el externalismo\*

*Nicolás Alles*

¿Qué nos hace actuar moralmente? ¿Qué instancias intervienen? ¿Qué sucede cuando esa motivación falla? Estos son algunos de los interrogantes implicados en el razonamiento moral y que han sido objeto de largas indagaciones filosóficas desde los orígenes mismos de la ética. Y aunque la preocupación central de las distintas teorías éticas parece ser la de la justificación de la obligación moral, la pregunta, con aristas más conectadas con la práctica, acerca de cómo es que nos motivamos a actuar de acuerdo con esa obligación no ha perdido un ápice de importancia. En este capítulo me encargaré de analizar una parte de esa gran discusión en torno a la motivación moral.

En aras de la claridad conviene comenzar por delimitar el ámbito del que me ocuparé. Si bien es posible encontrar en todo autor clásico de la ética normativa referencias al problema de la motivación moral, este asunto ha tenido particular relevancia en el ámbito de la metaética. Esta disciplina, que podríamos entenderla como aquella que se encarga de pensar los fundamentos del pensamiento moral normativo, ha hecho de la motivación moral uno de sus objetos de análisis más importantes. En este sentido, al hablar de motivación estamos refiriéndonos a un campo de reflexión particular de la filosofía práctica contemporánea que se caracteriza por una amplísima discusión que se mantiene hasta el presente.

Por todo lo anterior, y como repetiré varias veces a lo largo de las páginas que siguen, será necesario un cierto recorte para poder asir un asunto que se expande en ramificaciones cada vez más específicas y particulares. Mi objetivo aquí es brindar un panorama, una visión comprehensiva de un debate particular dentro del ámbito de la motivación moral, aquel que puede encontrarse entre la posición internalista y externalista. Como cualquiera que se interese en esta temática podrá notar a poco de recorrer la literatura especializada, que una reseña más o menos detallada de todas las implicancias y opciones plantea esta polémica consumiría mucho más del espacio que tengo disponible. En ese sentido, aquí solo propongo un mapa o estado de la cuestión que pueda servir de guía para futuras indagaciones. Por lo tanto, en lo que sigue, me limitaré a reconstruir las tesis principales de estas dos perspectivas sobre la motivación moral y a plantear algunos puntos de encuentro y desacuerdo entre estos enfoques.

---

\* Todas las citas de obras en inglés son traducciones propias.



Como intentaré mostrar, la diferencia central entre el internalismo y el externalismo concierne a la forma en la que se da la motivación moral, es decir, a la *modalidad* de esta. Más precisamente, estas dos concepciones implican dos maneras contrapuestas de entender cómo se da este tipo normativo de motivación: de manera necesaria para los internalistas o contingente para los externalistas. Esta controversia adquiere entonces la forma de dos puntos de vista que se excluyen mutuamente. Por eso, en las páginas que siguen, intentaré reproducir esta estructura dialógica entre estas dos opciones. Comenzaré por reconstruir las tesis centrales del internalismo para luego mostrar cómo las iniciativas externalistas parecen erigirse sobre la negación de esas tesis y contribuyen a la formulación de formas diferentes de entender la motivación.

¿Cuál es la relevancia contemporánea de atender a este debate? Tal como lo veo existen dos razones para visitar esta temática. En primer lugar, resulta una ocasión más que interesante para reevaluar el rol de la racionalidad en el razonamiento moral. Detrás de esta discusión entre la modalidad de la motivación moral pueden verse tesis sustantivas sobre qué idea hay de racionalidad y su función práctica en la moralidad. En segundo lugar, hay razones estrictamente prácticas para volver sobre este debate. El problema de la motivación moral no deja de ser una expresión más de la relación entre la teoría y la práctica. ¿De qué *forma* las exigencias morales se reflejan en la acción? Esta y otras preguntas marcaron la reflexión filosófica sobre la acción moral y encuentran en este debate (como en otros) un espacio especial de formulación y tratamiento.

El presente capítulo se ordenará de la siguiente manera. En primer lugar (2), me ocuparé de precisar los conceptos centrales que informaran el debate posterior entre el internalismo y el externalismo, me refiero a los conceptos de creencia, deseo y juicio moral. En segundo lugar (3), describiré las tesis y variantes principales del internalismo. En tercer lugar (4), describiré los argumentos más importantes del externalismo y cómo estos responden a las objeciones internalistas. Por último (5), concluiré con algunas reflexiones generales a modo de conclusión.

## **Elementos de la motivación moral.**

### **Deseos, creencias y juicio moral**

Antes de abordar qué es la motivación moral conviene dar cuenta de los elementos que intervienen en esta. El debate sobre esta clase de motivación es complejo, amplio y abarca diferentes aspectos y puntos de vista. No existe una sola teoría de la motivación moral, sino diversas perspectivas que atien-

den a los distintos aspectos involucrados, situación que nos lleva a realizar una serie de recortes necesarios para abordar nuestro tema. Me dedicaré en este apartado a precisar los conceptos más relevantes en torno a los cuales gira la cuestión de la motivación moral. Me refiero particularmente a los conceptos de deseo, creencias y juicio moral.

¿Qué nos mueve a actuar? Una visión de sentido común y que también es sostenida por varios filósofos afirma que si realizamos alguna acción, lo hacemos exclusivamente a causa de un deseo. Si alguien, pongamos por caso, asiste a escuchar un concierto para piano, lo hace porque tiene el deseo de escuchar la obra de algún compositor en particular o a algún pianista en especial. Apelar al deseo es la manera más directa que tenemos para explicar por qué alguien realiza (o no) alguna acción en particular. Sin embargo, existe otra forma de concebir a la motivación que quita protagonismo al deseo en el proceso de provocar la acción. Otros filósofos son de la opinión según la cual nuestra motivación a actuar está determinada por la creencia y no por el deseo. Alguien puede tener el deseo de salir a caminar por la playa, pero no lo hace porque sabe (tiene la creencia) que el sol, a esa hora del día, dañará su piel. Estas dos posiciones representan, como puede apreciarse, actitudes opuestas con respecto a la motivación en general y, como veremos, a la motivación moral en particular.

Deseo y creencia son dos tipos diferentes de estados mentales. Es posible diferenciarlos apelando a lo que se conoce como *dirección de ajuste*. Tanto el deseo como la creencia tienen una diferente relación con el mundo al que se *ajustan* de diferente manera. La creencia se ajusta al mundo; esto es, es el mundo el que dicta cómo debe ser la creencia para que esta pueda ser considerada como correcta o verdadera. En otras palabras, la creencia refleja al mundo ajustándose a lo que este es en un determinado momento. El deseo tiene, por su parte, una dirección inversa de ajuste con el mundo: de lo que se trata, en el caso del deseo, es que el mundo se ajuste al deseo. Aquí no se puede afirmar que el deseo sea verdadero o falso, el deseo no refleja al mundo, sino que el deseo es el que dicta cómo debe ser el mundo para que aquel se vea satisfecho.

Ahora bien, ¿cómo afecta esta distinción a la motivación moral en particular? En el ámbito de la motivación moral la diferencia entre el deseo y la creencia aparece de una manera evidente y marca la distancia entre dos formas tradicionales de comprender cómo nos motivamos a actuar moralmente. Al principio de este apartado comenté que existían distintas teorías de la motivación, dentro de esas existen un conjunto de ellas que tienen en cuenta justamente la importancia de los deseos y las creencias en la motivación; estas son conocidas como la teoría humeana de la motivación y la teoría anti-

humana de la motivación. Como puede sospecharse desde el nombre, estas teorías están inspiradas en aspectos de la filosofía moral del propio David Hume. Según la teoría humeana, «la motivación tiene su fuente en la presencia de un deseo relevante y una creencia medio-fin» (Smith, 2015:122). Es decir, no es posible pensar a la motivación moral de una manera independiente del deseo: sin deseo no hay motivación moral. Y, por lo tanto, la creencia por sí misma no es suficiente, aunque sí necesaria. Por su parte, la teoría anti-humeana sostiene lo contrario; la motivación moral no solo es posible sin la participación del deseo, sino que esta es posible apelando exclusivamente a la creencia o juicio moral.<sup>2</sup> La contraposición entre el humanismo y el antihumanismo en lo tocante a la motivación moral gira alrededor de determinar cuáles son los elementos necesarios y suficientes para lograr que alguien se encuentre motivado a actuar moralmente.

A la luz de estas distinciones es posible entender bajo una nueva perspectiva la diferencia entre el deseo y la creencia como dos estados mentales diferentes que implican direcciones de ajustes opuestas. Si atendemos a la definición que presenté más arriba de estos dos estados mentales parecería plausible afirmar que la creencia, por su propia definición, no podría motivar a actuar porque su función no es modificar el mundo de acuerdo con el criterio que dicta el deseo, sino representarlo. El antihumano debería entonces probar que la creencia puede motivar a actuar por sí sola. De hecho, existen versiones matizadas del antihumanismo que consideran que la creencia no solo tiene funciones representacionales, sino también motivacionales y que la imposibilidad de mover a la acción no sería tan inviable como piensa el humano.

En cualquier caso, la comparación entre estos dos abordajes parece actualizar en términos contemporáneos el debate moderno sobre la posibilidad misma de la razón práctica; de alguna manera, parecen actualizarse las diferencias entre David Hume e Immanuel Kant en torno a los alcances y límites de la razón en el ámbito práctico. En la actualidad, representantes de reformulaciones de estas teorías serían Bernard Williams, quien presenta un tipo particular de humanismo, o Thomas Nagel, cuyo antihumanismo recuerda al racionalismo de corte kantiano.

El último elemento que consideraré en esta sección es aquel referido al juicio moral. Todas las teorías de la motivación plantean una forma particular de relación con el juicio moral. En función de esto, debemos preguntarnos ahora ¿en qué consiste un juicio moral? O más precisamente aún, ¿qué es un juicio moral? En nuestras prácticas comunicativas existen diversos tipos de

---

<sup>2</sup> Thomas Nagel, en *La posibilidad del altruismo*, desarrolla una posición antihumana y sostiene que Hume hace descansar la ética, al incorporar la dimensión del deseo y la simpatía, en una contingencia psicológica (Nagel, 2004:21-23).

juicios que, por ejemplo, enuncian estados de cosas que podemos verificar en el mundo; podemos apelar a determinados juicios como por ejemplo «la luz está formada por partículas elementales llamadas fotones» o «tengo un número impar de libros en mi biblioteca» para describir estados de cosas constatables. Se trata, en estos casos, de juicios que refieren a situaciones fácticas. Sin embargo, los juicios morales tienen una naturaleza diferente. Como sostiene Valerie Tiberius (2015:75-76), la idea básica de los juicios morales es que estos nos mueven de una manera diferente a lo que pueden hacerlo los tipos de juicio mencionados. Esta idea con respecto a los juicios morales y su propiedad motivacional estaba presente, de nuevo, ya en Hume.

Los juicios morales entonces nos *disponen* a la acción. Estos tienden a hacernos actuar, pero esta tendencia puede ser seguramente frustrada por algunas circunstancias o superada por otras motivaciones. Este es el tipo de conexión especial que Hume piensa que tienen los juicios morales: no siempre nos motivan, pero están esencialmente asociados a la motivación. (...) Debido a que Hume piensa que los juicios hechos por nuestra facultad de razonar no son esencialmente motivacionales (no nos disponen necesariamente a la acción), concluye que los juicios morales deben provenir de nuestros sentimientos o pasiones, en vez de nuestra Razón. (Tiberius, 2015:76 – cursivas en el original)

Hasta ahora podemos afirmar que los juicios morales pueden definirse por una relación particular con la acción. Sin embargo, este aspecto no alcanza para definirlos de manera exhaustiva; decir que los juicios morales son motivacionales es pensarlos exclusivamente desde los efectos posibles que pueden generar y no por el contenido de los mismos. Existen otros juicios diferentes que también están ligados a la acción y no son morales como, por ejemplo, las normas legales o convencionales. Es a partir de esta distinción, entre las normas convencionales y las morales, que nuevamente Tiberius define el contenido de los juicios morales como juicios específicos que tienen una relación con la acción:

Se piensa que las normas morales deben ser más serias y tener una aplicabilidad más amplia que las convencionales. Las normas convencionales son consideradas como dependientes de una autoridad (como un maestro, la ley, ... o la cultura) y reciben un tipo particular de justificación diferente al de las normas morales, las cuales son justificadas usualmente en términos de daño o justicia. (2015:80-81)

Ahora es posible definir mejor los juicios morales; estos deben tener dos características: i) un contenido moral particular referido a la dimensión del daño y la justicia (lo cual, como repara Tiberius, implica un tipo de justifica-

ción especial); y ii) una determinada relación con la acción. Este último elemento será importante al definir las dos teorías sobre la motivación moral de las que me ocuparé en los apartados siguientes, el internalismo y el externalismo. Hasta aquí, la relación de los juicios morales con la acción —inspirada en Hume— implica una *disposición* a la acción, la cual puede adoptar, como mostraré más adelante, dos modalidades: puede tratarse de un vínculo *contingente* o de un vínculo *necesario* entre el juicio y la acción. El primer tipo de vínculo, el contingente, representa solo una perspectiva en la motivación moral, la externalista; sin embargo, el otro tipo de vínculo, el necesario, representa al internalismo que supone una relación de necesidad entre el juicio y la motivación. Ambas posiciones suponen entender a la motivación moral de maneras diferentes y presentan puntos atendibles y limitaciones claras para pensar el fenómeno de la motivación moral. Las dos próximas secciones estarán dedicadas a presentar estas dos maneras de pensar la motivación moral.

## **Internalismo**

En esta sección precisaré las características centrales del internalismo. Se trata de una de las formulaciones más discutidas en el ámbito de la metaética y con algunos adherentes que se cuentan entre los más importantes filósofos que estudian el problema de la motivación. El centro de la postura internalista, que intentaré explicitar en lo que sigue, apunta a proponer una conexión conceptual y necesaria entre las exigencias de la moralidad y la motivación de los agentes. Como aclararé a continuación, las exigencias de la moral pueden representarse de diferentes maneras, desde perspectivas realistas que hablan de hechos morales hasta los juicios morales a los que pueden adscribir los agentes. En cualquier caso, esta manera de interpretar la motivación moral sostiene que esas exigencias mueven a la acción sin ninguna mediación; solo basta con reconocer esas exigencias para que el agente se encuentre motivado. La conexión, como mencioné más arriba, es necesaria. Esta posición parece ciertamente intuitiva. ¿Qué otra cosa puede movernos a actuar si no son nuestras consideraciones morales? ¿Qué otra cosa puede llevar a cambiar la acción de un agente si no es una nueva evaluación moral? El internalismo ilustra estas intuiciones.

Los debates en metaética, y en especial aquel sobre el tópico de la motivación, se caracterizan por ser complejos, ricos y suelen ahondar en infinidad de variantes y asuntos específicos. El caso del internalismo no es la excepción. Basta revisar, aunque más no sea someramente, la literatura disponible sobre el internalismo para confirmar esta intuición. En la actualidad existen diferentes versiones y formas de entender la tesis internalista que brevemente esbozamos

más arriba. En consecuencia, hablar de esta forma de entender la motivación moral implica necesariamente operar sobre algún recorte de las posibles opciones disponibles de interpretarla. Aquí me ocuparé de describir tres variantes del internalismo que resultan ser las más discutidas en los debates contemporáneos. Si bien todas tienen puntos en contacto, pues parten de esta relación de necesidad entre la moralidad y la motivación, difieren en la forma de entender ese vínculo. Me refiero en particular al internalismo de existencia, al internalismo de razones y, por último, al internalismo de juicio moral.<sup>3</sup>

### Internalismo de existencia

De todas las clasificaciones que es posible encontrar sobre las variantes del internalismo, la propuesta por Stephen Darwall es tenida en cuenta por varios investigadores como punto de partida de la discusión.<sup>4</sup> Este autor propone una forma de internalismo al que denomina *internalismo de existencia*, el cual puede definirse de la siguiente manera:

Una forma diferente de internalismo es la tesis de que algo es el fundamento de un acto al tener cierta propiedad (ser correcto, por ejemplo, o tener el respaldo de razones) sólo si es capaz de motivar al agente. Esta clase de internalismo ubica un límite en la *existencia* de fundamentos para la corrección o racionalidad de un acto, y, por lo tanto, para que el acto pueda ser ya sea racional o correcto. En otras palabras, expresa una condición necesaria para la consideración de una razón o de un hecho que hace correcto [al acto]. Llamaremos a esta posición *internalismo de existencia*. (Darwall, 1983:54)

En esta forma de internalismo, las exigencias de la moralidad vienen dadas por la existencia de hechos morales o propiedades que hacen correcto a una acción o situación. De esta manera, este internalismo entiende que el elemento motivacional es el reconocimiento de ese hecho particular que convierte al acto en bueno. Como es posible apreciar, esta variante de internalismo no solo implica una tesis motivacional, sino también ontológica: la motivación descansa en alguna forma de realismo moral. Es por eso que Connie Rosati, en su entrada sobre la motivación moral para la *Stanford Encyclopedia of Philosophy*, propone como ejemplo de este tipo particular de internalismo a

---

<sup>3</sup> En la denominación de estas variantes del internalismo sigo la clasificación propuesta por Rosati (2016).

<sup>4</sup> G.F. Schueler (2010:294), por ejemplo, la toma como punto de partida de su propia clasificación.

la lectura que J.L. Mackie realiza de la teoría de las Ideas de Platón (Rosati, 2016 – sec. 3.2.). Efectivamente para Mackie:

la más importante tradición europea de filosofía moral, de Platón en adelante, ha combinado la idea de que los valores morales son objetivos con el reconocimiento de que los juicios morales son en parte prescriptivos o directivos o actúan como pautas de la conducta. Los propios valores han sido considerados como elementos simultáneamente prescriptivos y objetivos. En la teoría platónica, las formas, y especialmente la forma del bien, son realidades eternas, exteriores a la propia mente. Son un elemento estructural absolutamente central en la arquitectura del mundo, pero se afirma también que el simple hecho de conocerlas o «verlas» no se limitará a dictar a los hombres lo que deben hacer, sino que será garantía de que lo hacen, sobreponiéndose a cualquier inclinación contraria. (2000:26)

Esta forma de internalismo de existencia que Mackie encuentra en Platón implica una forma de realismo moral basado en una dimensión ontológica, las Ideas. En el proyecto moral (y ontológico) de Platón sólo basta con *ver* la Idea del Bien para estar necesariamente motivado a actuar de acuerdo con ella. Mackie sostiene que este tipo de argumentos aúna tres dimensiones: una ontológica, una racional y una motivacional que también puede encontrarse en otros autores clásicos de la tradición filosófica como Kant:

Según Platón, los reyes filósofos de la República merecen una confianza no contrapesada por ningún otro poder porque su educación les ha proporcionado el conocimiento de las ideas. Estando familiarizados con las ideas del bien, la justicia, la belleza y las demás ideas, se verán impulsados, por el mero hecho de su conocimiento y sin necesidad de motivación añadida alguna a ir tras esos ideales y a promoverlos. De manera similar, Kant cree que la razón pura puede ser práctica por sí misma, aunque no pretende ser capaz de explicar cómo es eso posible. (2000:26–27)

Esta apelación a Kant resulta pertinente para describir otra característica con la que suele asociarse al internalismo: el racionalismo. Tanto en Platón como en Kant, como antecedentes del internalismo, puede encontrarse la idea de que la mera razón, al reconocer el deber moral, debe encontrarse motivada a actuar de la manera en la que el deber lo indica sobreponiéndose a cualquier otra instancia motivacional que pueda aparecer, como los deseos o las inclinaciones. Actuar de manera contraria a lo que la razón identifica como deber sería, sobre todo para Kant, una forma de irracionalidad práctica. Esta concepción que considera a la racionalidad como central en la em-

presa de la motivación será retomada (de forma más matizada, claro) por filósofos que se reconocen abiertamente como internalistas, como el caso de Michael Smith al que volveremos más adelante.

### Internalismo de razones

La segunda variante de esta perspectiva es el *internalismo de razones*. Según este tipo de internalismo, alguien estará necesariamente motivado a actuar si tiene razones para actuar. Este tipo de internalismo es asociado con la argumentación desarrollada por Bernard Williams en su famoso artículo «Internal and external reasons». Allí sostiene una distinción que marcará el debate posterior concerniente a las razones para actuar: Williams distingue entre razones internas y razones externas para la acción. El punto por el cual se lo considera internalista es que argumentará que solo las razones internas pueden motivar y que las razones externas no son, en última instancia, posibles.

En el mencionado artículo, Williams, al explicar la interpretación interna de las razones, sostiene que «*A* tiene una razón para  $\varphi$  si *A* tiene algún deseo que se verá satisfecho si *A* hace  $\varphi$ » (1993:132). Los parecidos con el modelo de motivación propuesto por Hume son evidentes; sin embargo, Williams se encarga de aclarar que los argumentos de Hume son más complejos por lo que decide llamar a su propio modelo *subhumeano* (132). El punto central del internalismo de Williams supone la existencia de un *conjunto motivacional subjetivo* que todo agente posee, al que Williams denomina *S*, el cual está compuesto no solo de deseos, sino también de «disposiciones de evaluación, patrones de reacción emotiva, lealtades personales, y diversos proyectos, como podría llamárseles en forma abstracta, que incluyen compromisos del agente» (1993:136).

De esta manera, este modelo subhumeano de Williams afirma «que cualquier elemento de *S* origina una razón interna» (1993:132). En otras palabras, la motivación a actuar depende de que el agente tenga en su conjunto subjetivo motivacional un elemento relativo a la acción en cuestión que funciona como una razón interna. De esta manera, si solo lo que motiva la acción son las razones internas resulta, en consecuencia, que las razones externas —las que no pueden encontrarse en *S*— son simplemente imposibles. Williams lo explica con el caso del personaje de Owen Wingrave de Henry James. Según la historia, Owen es presionado por su familia para que se una al ejército sobre la base de razones basadas en el orgullo familiar y la tradición. Williams analiza esta historia de la siguiente manera:



Owen Winegrave no tiene motivación alguna para unirse al ejército, y todos sus deseos lo llevan en dirección distinta: odia la vida militar y lo que significa. Su familia podría haberse expresado diciendo que existe una razón para que Owen se una al ejército. Saber que no hay nada en el *S* de Owen que pudiera conducirlo a hacerlo, a través de la razón deliberativa, no los haría retirar la afirmación o admitir que la expresan bajo una interpretación errónea. El sentido con que la expresan es externo. (1993:137 – cursivas en el original)

Sin embargo, Williams sostiene que las razones externas no pueden ser motivacionales. Y si lo son es porque ya pertenecen a *S* y, por lo tanto, no son externas, sino internas. En palabras del propio Williams:

La idea de los enunciados sobre razones externas es que pueden ser verdad independiente de las motivaciones del agente. Pero nada puede explicar las acciones (intencionales) de un agente excepto que lo motive a actuar así. De manera que se necesita algo más aparte de la verdad del enunciado sobre razones externas para explicar la acción, algún enlace psicológico; y tal enlace psicológico parecen integrarlo las creencias. El hecho de que *A* crea un enunciado sobre razones externas sobre sí mismo puede ayudar a explicar su acción. (1993:138 – cursivas en el original)

Por último, agrega: «Una persona que cree que las consideraciones de honor familiar constituyen razones para actuar tiene cierta disposición para actuar, y también disposiciones de aprobación, sentimiento, reacción emocional, etcétera» (1993:139).

Si las razones externas son en realidad «disposiciones de aprobación, sentimiento, reacción emocional», entonces pertenecen al *S* del agente y por lo tanto solo parecen ser razones externas, pero en realidad son internas. De acuerdo con esta interpretación, las razones externas serían imposibles; su potencial motivacional solo revelaría que en realidad son internas. De esta manera, para Williams solo es posible una forma internalista de motivación, pero no como en el caso que Mackie refiere a Platón y a Kant, basado en hechos morales; en este caso, el internalismo sería de razones que tiene el sujeto para motivarse a actuar.

## Internalismo de juicio moral

Ahora me ocuparé de describir la tercera variante del internalismo, la cual es el centro de discusiones en el debate contemporáneo sobre motivación moral; me refiero al *internalismo de juicio moral*. De acuerdo con esta concepción del internalismo, los juicios morales son intrínsecamente motivacionales, esto es, motivan sin la necesidad de la colaboración de ningún estado conativo particular (Svavarsdóttir, 1999:163). Como aclara Sigrún Svavarsdóttir, esta forma de internalismo implica una necesidad conceptual en relación con el contenido de los juicios morales: «los juicios morales están necesariamente conectados con la motivación a perseguir o promover aquello que es juzgado favorable y a evitar o prevenir aquello que es juzgado como desfavorable» (1999:163). Desde este punto de vista un agente no podría decirse que realmente haya hecho un juicio moral a menos que se encuentre motivado a actuar de acuerdo con él.

El problema con este internalismo es, como puede imaginarse, la carga que supone este requisito de la necesidad conceptual que implica. En otras palabras, esta comprensión del internalismo pasa por ser extremadamente exigente y no tiene en cuenta la serie de factores que pueden intervenir para que alguien, incluso si cree sinceramente en el juicio moral que expresa, pueda no encontrarse motivado a actuar de acuerdo con ese juicio. En las discusiones de metaética en torno a estos problemas se suele hablar de la debilidad de la voluntad como una situación que afecta a nuestra capacidad de actuar de acuerdo con lo que creemos que es nuestra obligación. Este es un fenómeno usual y se relaciona directamente con el internalismo.

La definición del internalismo de juicio moral que hicimos más arriba parece no tener en cuenta este tipo de fenómenos (al que habría que agregar también otros como la depresión) al momento de pensar la motivación moral y su relación tan supuestamente necesaria con el juicio moral. Advertidos de estas falencias algunos internalista proponen versiones un poco más matizadas del internalismo que vislumbrarían estos casos. Michael Smith, uno de los exponentes más importantes de esta corriente, aborda estas cuestiones en *El problema moral*. Allí sostiene que es posible formular una concepción del internalismo en la que, si bien existe una conexión conceptual entre el juicio moral y la motivación, esta conexión puede ser en ocasiones derrotada. Para dar cuenta de lo anterior, Smith propone tres maneras diferentes de entender la tesis internalista:

1. «Si un agente juzga que es correcto para él hacer  $\varphi$  en la circunstancia  $c$ , entonces está motivado a hacer  $\varphi$  en  $c$ » (2015:92).

2. «Si un agente juzga que es correcto para él hacer  $\phi$  en la circunstancia  $c$ , entonces o bien está motivado a hacer  $\phi$  en  $c$  o bien es prácticamente irracional» (92).
3. «Si es correcto que los agentes hagan  $\phi$  en las circunstancias  $c$ , entonces hay una razón para que estos agentes hagan  $\phi$  en  $c$ » (93).

Para Smith, la afirmación (1) es una lectura general del internalismo que pone de manifiesto la conexión conceptual entre el juicio moral y la voluntad. Por su parte, (3) es considerada por Smith como la afirmación racionalista del internalismo, implica a (2) y sostiene que los hechos morales son ellos mismos requerimientos de la racionalidad o de la razón (93). Por último, (2) parece ser un punto medio entre las otras dos posiciones, la cual es denominada «el requerimiento de carácter práctico sobre el juicio moral» y es la visión que caracterizará al internalismo (93). La ventaja que tiene (2) sobre (3) es que puede ser defendida por quienes niegan a (3), la afirmación racionalista. De ahí que la propia posición de Smith sea una versión débil del internalismo al disminuir los requerimientos fuertes de racionalidad que plantea (3).

La manera que encuentra Smith de defender su posición apunta a mostrar que (2) constituye la manera más *confiable* de explicar la relación entre el juicio moral y la motivación de la persona buena y de buena voluntad. Más precisamente, Smith se pregunta «¿Cómo hemos de explicar por qué, en cierto rango de circunstancias contrafácticas las motivaciones morales de la persona buena y de buena voluntad siempre estarán alineadas con sus nuevos juicios morales?» (2015:102). Smith lo explica de la siguiente manera:

Como yo lo veo, sólo hay dos respuestas posibles. Por un lado, podemos decir que la conexión confiable entre juicio y motivación tiene que ser explicada internamente: se sigue directamente del contenido del juicio moral mismo. La idea será entonces que o bien la creencia de que un acto es correcto produce la motivación correspondiente (ésta es la alternativa racionalista), o bien que la actitud de aceptar que un acto es correcto es ella misma idéntica con el estado de estar motivado (ésta es la alternativa expresivista). Por otro lado, podemos decir que la conexión confiable entre el juicio y la motivación debe ser explicada externamente: se sigue del contenido de las disposiciones motivacionales poseídas por la persona buena y de buena voluntad. Aquellos que defienden el requerimiento de carácter práctico optan por la primera respuesta, los externalistas fuertes optan por la segunda. (Smith, 2015:102)

En otras palabras, hay solo dos opciones para entender la relación entre el juicio moral y la motivación: la internalista y la externalista. Smith defenderá, por supuesto, la perspectiva internalista sobre la base de que es la única forma

de entender esta vinculación de manera no fetichista. Smith lo aclara recurriendo a la motivación típica de la buena persona:

Pues el sentido común nos dice que si la gente buena juzga correcto ser honesto, o cuidar a los hijos y a los amigos y a los afectos, o que juzga correcto que la gente reciba lo que se merece, entonces se preocupan no derivativamente sobre estas cuestiones. Las personas buenas se preocupan no derivativamente sobre la honestidad, sobre las idas y venidas de sus hijos y amigos, el bienestar de sus afectos, de que la gente obtenga lo que merece, la justicia, la igualdad y cosas por el estilo, y no se preocupan sobre una cosa: hacer lo que creen que es correcto cuando esto es leído *de dicto* y no *de re*. De hecho, el sentido común nos dice que estar motivado de ese modo es un fetiche o un vicio moral, no una virtud moral. (2015:104)

Aquí se plantean dos posibles variantes de entender el deseo de actuar correctamente: *de dicto* y *de re*. Si alguien se encuentra motivado a actuar por un deseo *de dicto* de hacer lo que considera correcto, su deseo implica la idea de corrección moral. En cambio, si la persona se encuentra motivada por un deseo de hacer lo correcto leído *de re*, entonces su deseo no contiene el concepto de corrección moral.<sup>5</sup> En este sentido, una persona buena y de buena voluntad desea hacer lo correcto leído *de dicto*; esto es: desea hacer el bien. En cambio, cuando una persona desea hacer cosas buenas *de re*, en realidad, desea hacer cosas que resultan buenas, pero no las desea *porque* sean buenas, sino que está deseando otra cosa que lo lleva a hacer cosas buenas.

Imaginemos que alguien desea hacer cosas buenas y su deseo es leído *de dicto*; en ese caso, ayudará a sus amigos porque se preocupa *intrínsecamente* por el bienestar de ellos. Pensemos ahora en alguien que desea hacer cosas buenas y su deseo ahora es leído *de re*; en este caso, en cambio, ayudará a sus amigos, pero de una manera derivada, esto es, su preocupación principal puede ser hacerse fama de bueno o impresionar a alguien y por eso desea ayudar a sus amigos. Entre estos dos casos hay diferencias en la valoración moral de la persona: solo en el primer caso podemos decir que se trata de una buena persona en los términos que plantea Smith. Para este autor, el internalista desea hacer lo correcto *de dicto*, y el externalista *de re*. En este sentido, el externalista tiene su deseo *fetichizado*: desea hacer las cosas buenas por *algo más*, distinto al deseo intrínseco de actuar correctamente.

Todo lo anterior es lo que le hace concluir a Smith que solo el internalista puede cumplir con el requerimiento de carácter práctico sobre el juicio moral

---

4 Para una explicación de la diferencia entre los deseos de dicto y de re, ver Strandberg (2000).

y alinear su motivación con los juicios morales. De esta manera, el internalista cambiará de motivación al cambiar sus juicios morales, algo que el externalista no podrá hacer, a juicio de Smith, porque el contenido de su motivación está *fetichizado*, es decir, no depende de sus juicios morales, sino de algo diferente a ellos. En conclusión, la única forma de cumplir con las exigencias morales es adoptar el internalismo, y en particular la formulación expresada en (2) que a su vez no implica una demanda de racionalidad tan grande como (3). En palabras de Smith:

Por tanto hemos encontrado una razón decisiva para rechazar la explicación externalista fuerte de la conexión confiable entre el juicio moral y la motivación en la persona buena y de buena voluntad. Pues, en resumen, la explicación del externalista fuerte nos compromete a visiones falsas sobre el contenido de las motivaciones de las buenas personas. Eleva a un fetiche moral como la única virtud moral. Y el remedio, por supuesto, es retirarse y adoptar la explicación alternativa, internalista, de la conexión confiable entre el juicio moral y la motivación. Pero si hacemos eso entonces, por supuesto, tenemos que aceptar el requerimiento de carácter práctico como una exigencia en el contenido del juicio moral después de todo. (2015:105–106)

Este planteo de Smith se suele conocer también como la objeción del fetiche moral dirigida al externalismo. Existen diferencias sobre los alcances y limitaciones de esta objeción. Sobre ella, y sobre las posibilidades de responderla, me ocuparé más adelante al tratar las tesis externalistas.

## **Externalismo**

La otra posición en el debate sobre la motivación moral es, como dijimos, el externalismo. Podríamos definirla —como de hecho se suele hacer— como la postura que niega la tesis central del internalismo. Más precisamente, el punto en el que internalismo y externalismo se oponen es el tipo de vínculo que existe entre el juicio moral y la motivación. Para el externalista no existiría una conexión necesaria entre la aceptación de un juicio moral y la motivación a actuar de acuerdo con lo que establece este juicio. En su reconstrucción del externalismo, David Brink le asigna tres características que se oponen abiertamente al internalismo. Según Brink, el externalista puede afirmar:

en primer lugar que las consideraciones morales solo motivan o justifican de manera *contingente*; en segundo lugar, que el poder motivacional o racionalidad de la moralidad, ya sea este necesario o contingente, solo puede conocerse *a posteriori*; o en tercer lugar, el poder motivacional o racionalidad de la moralidad, ya sea necesario o contingente, a priori o a posteriori, depende de cosas diferentes de nuestro concepto de moralidad, tales como el contenido de lo que moralidad resulta, una teoría sustantiva de las razones para la acción o los hechos acerca de los agentes como sus intereses o deseos. (Brink, 1989:42-43 – cursivas en el original)

El externalismo complejiza el panorama en torno a la motivación al incorporar elementos que no estaban presentes en la consideración del internalismo para lograr la motivación moral, entre ellos, los más importantes son los estados conativos (Svavarsdóttir, 1999:163). Los adherentes a la estrategia externalista sostienen que lo que logra efectivamente la motivación moral es un juicio moral en combinación con un estado conativo particular, un deseo.<sup>6</sup> Una de las formas de definirlo es la que propone Nick Zagwill cuando afirma: «El externalismo motivacional es la perspectiva según la cual los juicios morales no tienen eficacia motivacional en sí mismos, y que cuando motivan, la fuente de motivación reside fuera del juicio moral en un deseo separado» (2003:143).

### Elementos y alcances de la motivación externalista

Esta interpretación de la motivación reporta algunas ventajas sobre su contraparte internalista. Más precisamente, el externalismo permite explicar algunos casos de los que el internalismo no podía dar cuenta. ¿Cómo explicar los casos en los que los agentes podían reconocer perfectamente el hecho de que tenían una obligación moral, pero sin embargo no lograban estar motivados a actuar de acuerdo con ese mandato moral? O incluso, ¿cómo considerar al agente que decide deliberadamente actuar mal? De lo que se trata aquí es del problema que algunos filósofos, como David Brink, identificaron como el desafío del amoral. En resumen, este desafío supone la posibilidad de que haya alguien que frente a las exigencias expresadas en los juicios morales permanezca indiferente a ellas y no logre motivarse a actuar de acuerdo con esas exigencias.

---

<sup>6</sup> En este sentido, la posición externalista así definida tendría afinidades con el humeanismo descripta más arriba.

La respuesta del internalista frente a este desafío presenta dos variantes. En primer lugar, algunos internalistas sostienen que quien se encuentra en esta situación descripta recae en alguna forma de incapacidad por parte del agente. Más precisamente, se trataría de una incapacidad racional o, como sostiene Michael Smith, el amoral no sería *plenamente* racional.<sup>7</sup> En segundo lugar, la otra respuesta apunta a un uso indebido de los términos morales. El amoral no estaría usando los términos morales de manera adecuada al no verse motivado por ellos, sino que los estaría usando de manera entrecomillada, es decir, en un sentido figurado, por eso se dice que el amoral no está *sinceramente* comprometido con ese juicio moral incluso si dice aceptarlo (Brink, 1989:46). El estrecho margen de opciones que ofrece el internalismo para explicar esto sólo revela, tal como observa pertinentemente Sigrun Svavarsdóttir, que son vistos como una imposibilidad conceptual (1999:166). Simplemente resultan inconcebibles. Al consultar la experiencia queda claro que estas dos respuestas del internalista no son del todo satisfactorias.

No es difícil imaginar e incluso constatar casos en los que esa imposibilidad conceptual que prefigura el internalismo no es tal. Es posible verificar situaciones en que el agente adhiere sinceramente un juicio moral y, al mismo tiempo, es plenamente racional, pero, de todos modos, no se encuentra motivado a actuar de acuerdo con ese juicio. Imaginemos, por ejemplo, a alguien que decide mentir a un amigo para no tener que aceptar su invitación a cenar, sabiendo que mentir es un acto moralmente incorrecto.<sup>8</sup> ¿Es esto una imposibilidad conceptual como lo supone el internalismo? Ciertamente no. Por el contrario, el externalismo puede encontrar una explicación mejor que el internalismo para dar cuenta de estos casos al postular una conexión diferente entre las exigencias morales y la motivación. David Brink destaca tres formas en las que el externalismo puede explicar de manera más satisfactoria que su contraparte esta situación:

El externalismo proporciona una explicación más plausible de la conexión entre la moralidad y la motivación; hace de la fuerza motivacional de las consideraciones morales una cuestión psicológica contingente, dependiendo de

---

7 De hecho, la formulación del propio internalismo de Michael Smith expresada en «el requerimiento de carácter práctico sobre el juicio moral» implica que un agente que juzga correcto realizar una acción y no se encuentra motivado es «prácticamente irracional».

8 Entiendo que el concepto de amoralista que plantean las discusiones refiere a personas que no logran motivarse por los juicios morales de manera sistemática. En ese sentido, el concepto de amoralista abarcaría mucho más que el ejemplo reducido que brindo aquí. De todos modos, un amoralista podría perfectamente comportarse como el agente imaginario que propongo en este ejemplo. Por lo tanto, creo que situación imaginaria que sugiero, aunque local y circunscripta, no pierde valor explicativo.

las creencias y los deseos que tengan los agentes. En primer lugar, porque no hace de la fuerza motivacional de las obligaciones morales una característica conceptual de la moralidad, el externalismo no toma al alcance y al rigor de las obligaciones morales como rehenes de los deseos reales de las personas (como el internalismo del agente). En segundo lugar, el externalismo también hace que la fuerza motivacional del juicio moral y la creencia moral sea un hecho psicológico contingente, que dependa tanto del contenido de los puntos de vista morales de las personas como de sus actitudes y deseos. (...) Por último, el externalismo nos permite tomar en serio el amoralismo y el desafío amoralista. Dado que el externalismo no trata de encontrar la fuerza motivacional dentro de las propias consideraciones morales, reconoce que podemos imaginar a alguien que reconozca las consideraciones morales y permanece impasible. (1989:49)

La explicación externalista parece más plausible porque considera un hecho que el internalismo parece pasar por alto y que Brink rescata en este fragmento: el estado psicológico del agente que actúa. El internalismo, como vimos, al proponer una conexión necesaria entre las exigencias de la moral y la motivación, deja de lado cualquier instancia contingente del sujeto que actúa y supone que la moral tendrá una fuerza motivacional prioritaria por sobre todas las otras inclinaciones que ese sujeto pueda tener en un determinado momento. ¿Es plausible asumir como cierto este supuesto del internalismo? O, en otras palabras, ¿es posible asumir que *siempre y en todo caso* las razones morales motivarán de la misma manera? Tal vez, la respuesta más prudente sea decir que no. En este sentido, la tensión entre el internalismo y el externalismo con respecto al desafío del amoralista parece reproducir las tensiones entre los modelos racionalistas y aquellos que cuestionan los límites de estos. No resulta extraño que los planteos racionalistas (incluso los considerados como *débiles*) adhieran a alguna forma de internalismo,<sup>9</sup> mientras que los planteos externalistas se relacionen con los planteos como los *humeanos*.<sup>10</sup> De nuevo, Brink, puede echar luz sobre este asunto:

Sin embargo, el externalismo puede basar esta motivación en hechos psicológicos «profundos» o ampliamente compartidos. Pensemos en los puntos de vista morales más comunes que reconocen el carácter del respeto al otro de muchas exigencias morales. Si, por ejemplo, la simpatía es, como sostenía Hume, un rasgo psicológico profundamente arraigado y ampliamente compartido, entonces, como una cuestión de hecho psicológica contingente (pero «profunda»), la

---

9 Ejemplos de estos son los trabajos de Thomas Nagel y Michael Smith.

10 Aunque esto no quita que haya *humeanos* entre los internalistas, como vimos en el caso de Williams.



gran mayoría de las personas tendrán al menos *algún* deseo de cumplir con lo que perciben son sus obligaciones morales, incluso con aquellas obligaciones morales que tienen en cuenta a los otros. La motivación moral, desde esta perspectiva, puede ser ampliada y predecible, aunque no sea necesaria, ni universal, ni prioritaria. Estas son las limitaciones de la fuerza motivacional real de las consideraciones morales que, creo, la reflexión sobre la moral de sentido común reconoce. (1989:49 – cursivas en el original))

¿Alcanza lo anterior para responder al interrogante acerca de si debemos adoptar una concepción externalista de la motivación? Probablemente no. Existen variantes internalistas reformuladas que intentan asimilar su posición con respecto al amoralista y superar estos escollos.<sup>11</sup> Sin embargo, resulta claro que el debate sobre el amoralista es una ocasión para mostrar los límites y alcances de cada posición en pugna. El externalismo mostró apoyarse en una psicología moral más realista y compleja que aquella que propone la impronta racionalista de los modelos internalistas que deja de lado aspectos tales como las circunstancias del agente y el conjunto de deseos que este puede tener.

¿Puede el externalismo responder a las críticas internalistas?

Sin embargo, el externalismo tiene problemas propios. Si bien parece poder responder mejor que su contraparte a los casos en los que el reconocimiento de una obligación moral no logra motivar, aportando una versión más completa de los elementos que componen la motivación moral al incorporar al deseo, esta interpretación de la motivación moral presenta inconvenientes que les son inherentes y en algún sentido no menos importantes que los que pretendía resolver. En este apartado me ocuparé de dos objeciones clásicas al externalismo y exploraré la posibilidad que tiene esta forma de motivación de responder a ellas. Más precisamente, me refiero a i) la objeción del fetiche moral y ii) la objeción de los deseos egoístas. Trataré por turnos estas objeciones y analizaré las diferentes maneras de responderlas. Comenzaré por la primera, por la objeción del fetiche moral.

Cuando presenté el internalismo de juicio de Michael Smith mencioné que esta forma de internalismo, según Smith, era la única que podía explicar la confiabilidad entre el juicio moral y la motivación de la persona buena y de buena voluntad. En ese punto, Smith apunta a la distinción entre el deseo de

---

11. De hecho, la posición internalista desarrollada por Smith en *El problema moral* pretende responder explícitamente a Brink en este punto.

hacer lo correcto leído *de dicto* y leído *de re*. Como intenté señalar más arriba, una persona que desea hacer el bien y ese deseo es leído *de dicto* desea hacer las cosas que son buenas; por su parte, quien desea lo mismo pero ese deseo es leído *de re* en realidad no desea hacer las cosas porque son buenas, sino que las desea porque tienen ciertas características que las hacen buenas. La diferencia es que quien lee el deseo *de dicto* desea hacer las cosas *intrínsecamente* buenas y quien lee ese deseo *de re* desea hacer cosas buenas, pero de manera *derivada*, no porque desee intrínsecamente esas cosas buenas. Smith concluye que solo el internalismo puede explicar la relación entre el juicio moral y la motivación en términos de un deseo leído *de dicto* y, por lo tanto, el externalismo fetichiza su deseo de hacer las cosas correctas.

Tradicionalmente este argumento fue entendido como una objeción al externalismo, la objeción del fetiche moral. La implicancia central que tiene para el externalismo es que le niega la posibilidad de alinear el cambio de juicio moral con la motivación. Solo alguien que lee el deseo de hacer las cosas correctas *de dicto* puede, luego de cambiar su juicio moral, cambiar de forma análoga su motivación a actuar. El externalista, al leer ese mismo deseo *de re*, no puede explicar esa alineación y, por lo tanto, un cambio de juicio moral no implicaría —al menos desde la perspectiva internalista de Smith— un consiguiente cambio en su motivación. La pregunta que resta hacernos es ¿realmente el externalismo no puede explicar el cambio de alineación entre el juicio moral y su motivación? A continuación, reconstruiré una posición que difiere con los internalistas en este punto.

En su artículo, «Externalism and the Content of Moral Motivation», Caj Strandberg explora la posibilidad de responder a la objeción del fetiche moral dirigida al externalismo. De acuerdo con su propuesta es posible pensar una alineación entre el juicio moral y el contenido de la motivación para el caso de los externalistas (algo que, como intenté señalar era del todo imposible para Smith). Es más, Strandberg considera que ese cambio puede realizarse sin la necesidad de adoptar una lectura *de dicto* del deseo de hacer las cosas correctas. De esta manera, sería posible concebir a una persona buena y de buena voluntad a partir de la lectura *de re* del deseo de actuar correctamente. Su argumento procede por partes. En primer lugar, nos pide que imaginemos la siguiente situación.

Imaginemos una persona buena y de buena voluntad que inicialmente juzga que no es correcto dar dinero a los carenciados que están en el centro de la ciudad, pero cambia su perspectiva y llega a juzgar que es correcto darles dinero, y eso la hace cambiar concordantemente su motivación: antes del cambio ella no está motivada a dar dinero a ningún carenciado, pero después del cambio está motivada a hacerlo y su resistencia previa desapareció. (2000:253)

La intención de Strandberg es mostrar que es posible explicar este cambio de alineación entre el juicio y la motivación en el contexto del externalismo. La segunda parte de su argumento afirma lo siguiente:

Asumamos en primer lugar que la persona en cuestión cree que ciertas características hacen a la acción correcta. En aras de la simplicidad podemos llamar F a estas características. Para ilustrar esto, supongamos que F consiste, entre otras cosas, en ayudar a las personas que lo necesitan. Luego, supongamos que esta persona tiene el deseo de realizar acciones que tienen la característica F. Los externalistas pueden ahora dar cuenta de la conexión confiable en discusión al asumir que la persona en cuestión altera sus creencias en lo que concierne a las que acciones que contienen F. Antes del cambio del juicio moral, ella no cree que darle dinero a los carenciados tiene F y consecuentemente juzga que no es correcto darles dinero. Sin embargo, por alguna razón llega a creer que darles dinero a los carenciados tiene F. (...) Como resultado, después del cambio en el juicio moral cree que darle dinero a los carenciados tiene F y juzga en consecuencia que es correcto darles dinero. Ahora, ya que tiene el deseo de realizar acciones que tienen F, su motivación cambia en concordancia. Esto es, antes del cambio en el juicio moral la persona no está motivada a dar dinero a los carenciados, mientras que se encuentra motivada a hacerlo después del cambio y su resistencia ha desaparecido. (2000:253)

¿Qué nos dice este ejemplo que propone Strandberg? La confiabilidad de la alineación entre el juicio moral y la motivación no depende exclusivamente de adoptar una lectura *de dicto* del deseo de hacer las cosas correctamente. Esa alineación puede dar también cuando leemos al deseo *de re*. La persona del ejemplo no desea ayudar a los carenciados porque desea *intrínsecamente* su bien, desea realizar algo en particular, algo que tiene la propiedad F, ese es su objetivo primordial. En el ejemplo resulta que al realizar acciones que tienen la propiedad F está ayudando indirectamente a los carenciados, es decir, su deseo de ayudarlos es derivado, no intrínseco. Sin embargo, ese cambio de motivación operó una vez que su juicio moral cambió: pasó de no desear realizar las acciones que contengan F a desear las acciones que contienen F que, recordemos, significa en este caso ayudar a las otras personas. En otras palabras, es posible desear ayudar a los demás sin adherir a una lectura *de dicto* y promoviendo una lectura *de re*. En conclusión, es factible ser una persona buena y de buena voluntad adoptando una forma de motivación externalista. En términos del autor que estamos analizando:

Como se observa, el deseo que figura en este tipo de explicación es un deseo *de re* de hacer lo que se juzga correcto. En el ejemplo mencionado es el deseo de realizar acciones que tienen cierta característica F que consiste, entre otras cosas, ayudar a aquellos que están necesitados. De acuerdo con este tipo de explicación, no hay necesidad de asumir que las personas buenas y de buena voluntad están motivadas por un deseo *de dicto* para explicar la conexión confiable. (2000:253 – cursivas en el original))

Ahora me ocuparé de la segunda objeción que mencioné más arriba, la objeción del deseo egoísta. De acuerdo con esta crítica, el externalismo disolvería la moralidad misma al hacer depender la motivación de un deseo externo. El núcleo de esta objeción afirmaría que si hacemos depender la moralidad de los deseos, como supone el externalismo, hacemos depender la moralidad de algo no moral. Nick Zangwill la expone sucintamente cuando dice:

Si las acciones son siempre el resultado de deseos en conjunción con creencias, entonces todas las acciones están motivadas por consideraciones egoístas. Pero las motivaciones morales y las egoístas son diametralmente opuestas. Por lo tanto, la fuerza motivacional de los juicios morales no puede ser externa a ellos. (2003:145)

Surgen así las siguientes preguntas: ¿son todos los deseos necesariamente egoístas? ¿Existe la posibilidad de pensar deseos morales? Los intentos de respuestas a esta objeción exploran esta posibilidad y sugieren que la dimensión conativa no implica de manera necesaria que el agente tenga solo deseos egoístas (Zangwill, 2003:145). En otras palabras, es posible que existan deseos morales y por lo tanto no egoístas. Quien sostendrá esto es Sigrún Svavarsdóttir en su artículo «Moral cognitivism and motivation». En sus propias palabras:

La disposición a estar motivada por los propios juicios morales se basa en una actitud conativa (deseo) adoptada hacia los objetos bajo una forma de presentación moral. Por comodidad, llámese deseo de ser moral. Supongo que varía de un agente a otro en los que otros estados mentales sostienen ese deseo. En algunos casos, puede estar sostenido por una buena dosis de simpatía hacia los demás, así como por la aceptación de una norma de benevolencia. En otros casos, en cambio, o adicionalmente, estará sostenido por el deseo de poder justificar la propia conducta desde un punto de vista imparcial. Pero, sin duda, también hay casos en los que se sustenta, por ejemplo, en el temor a Dios y en una visión teológica de los fundamentos de la moral, o en el miedo al castigo. También es posible que se sostenga por sí solo, sin depender de ningún otro motivo en esa etapa concreta

de la vida de la persona. Dudo que haya una psicología única que sostenga una disposición a estar motivada por los propios juicios morales. (1999:170)

Esta concepción de la motivación moral según Svavarsdóttir supone una comprensión del pensamiento moral que apela a recursos empleados en la formación de estados cognitivos, conativos y emocionales. Esta concepción, que desarrolla Svavarsdóttir y que otros filósofos externalistas comparten,<sup>12</sup> resulta una interesante reformulación de las tesis externalistas por dos motivos.

En primer lugar, constituye una concepción más amplia que la del racionalismo asociado a los planteos internalistas al reconocer que en la formación del pensamiento moral la dimensión cognitiva es solo un elemento más, al incorporar elementos conativos y emocionales. En segundo lugar, rebate la objeción de que los deseos solo tienen un objeto egoísta. Svavarsdóttir sostiene justamente que los deseos pueden tener diversos contenidos y perseguir diferentes objetos sin que esto los haga necesariamente egoístas o desprovistos de un contenido específicamente moral. Entre los ejemplos que sugiere Svavarsdóttir menciona que es posible advertir que el deseo de ser moral se sostiene sobre la base de, por ejemplo, la simpatía o al temor de Dios. En cualquier caso, Svavarsdóttir está preocupada por la existencia misma del deseo de ser moral y no por la fuente de ese deseo.

De lo que se trata en el fondo (e íntimamente relacionado con la objeción anterior) es reconocer que el internalismo no es la única forma de responder a las exigencias morales y eso trae de nuevo a colación la dimensión de lo que implica ser una persona buena y de buena voluntad. En la perspectiva externalista de Svavarsdóttir (1999:98), «una buena persona es también considerada, compasiva, amable, leal y honesta. Y estas características no implican disposiciones motivacionales que están comprometidas por representaciones morales distintivas de las propias circunstancias o las alternativas comportamentales» como pueden derivarse del modelo internalista. En otras palabras, los externalistas «están comprometidos con sostener que el deseo de ser moral es *una parte* de la estructura motivacional de la buena persona» (199 – cursivas en el original).

La posibilidad de considerar a estas respuestas como correctas señala que el debate entre el internalismo y el externalismo se encuentra lejos de cerrarse.

---

12 Aunque menos desarrollada, Nick Zangwill también sostiene la idea de un externalismo basado en la idea de un deseo de ser moral.

## Conclusión

En las páginas precedentes intenté reconstruir las principales coordenadas del debate entre el internalismo y el externalismo en el ámbito de la motivación moral. Como mencioné en la introducción, mi objetivo fue, sobre todo, brindar una visión panorámica de los principales elementos y discusiones que es posible encontrar en la contraposición de estos dos modelos de motivación moral. Tal como intenté exponer, este debate puede leerse como un juego de desafíos y respuestas sobre la naturaleza de la motivación moral. Cada tesis internalista es discutida y contestada por alguna variante externalista, cuestión que reinicia la controversia. Aquí me limité a algunas cuestiones centrales que no agotan el panorama completo que comprende esta discusión. En cualquier caso, entiendo que el valor de este asunto reside en sus aportes al momento de pensar las dimensiones, las características, los límites y los alcances de la racionalidad práctica y la posibilidad de ampliarla, cuando resulte muy estrecha, por elementos que también son esenciales en la dimensión práctica como los deseos u otros estados psicológicos.

## Referencias bibliográficas

- Brink, D. (1989). *Moral realism and the foundations of ethics*. Cambridge University Press.
- Darwall, S. (1983). *Impartial reason*. Cornell University Press.
- Mackie, J.L. (2000). *Ética. La invención de lo bueno y lo malo*. Gedisa.
- Nagel, T. (2004). *La posibilidad del altruismo*. Fondo de Cultura Económica.
- Rosati, C. (2016). Moral Motivation. En Zalta, E. (Ed.), *The Stanford Encyclopedia of Philosophy*. The Metaphysics Research Lab. Philosophy Department. Stanford University. <https://plato.stanford.edu/archives/win2016/entries/moral-motivation>
- Schueler, G.F. (2010). Motivational Internalism and Externalism. En O'Connor, T. y Sandis, C. (Eds), *A Companion to the Philosophy of Action* (293–300). Blackwell.
- Smith, M. (2015). *El problema moral*. Marcial Pons.
- Strandberg, C. (2000). Externalism and the Content of Moral Motivation. *Philosophia*, 35, 249–260.
- Svavarsdóttir, S. (1999). Moral Cognitivism and Motivation. *The Philosophical Review*, 108(2), 161–219.
- Tiberius, V. (2015). *Moral Psychology. A Contemporary Introduction*. Routledge.
- Williams, B. (1993). Razones internas y externas. En Williams, B., *La fortuna moral* (131–145). Universidad Nacional Autónoma de México.
- Zangwill, N. (2003). Externalist Moral Motivation. *American Philosophical Quarterly*, 40(2), 143–154.

## 18. *Akrasia* o debilidad de la voluntad

Luciana Samamé

Las controversias filosóficas sobre el fenómeno de la *akrasia* poseen un largo e intrincado historial. La discusión se inaugura con el *Protágoras* de Platón, para ser luego revistada por Aristóteles, los estoicos, el cristianismo de los primeros siglos y también el de la escolástica, por solo nombrar importantes hitos. En la Modernidad no pareció recibir atención privilegiada y, de hecho, recién promediando la mitad del siglo xx comenzó a suscitar renovado interés principalmente de parte de la filosofía analítica.<sup>1</sup> Este recorrido vasto y abigarrado nos permite caracterizar al debate sobre la *akrasia* como a uno de los más prolíficos que cabe encontrar entre las discusiones clásicas de la filosofía. Pero también, como a uno de los más difíciles de asir en términos claros y distintos —para emplear el lenguaje cartesiano—. Pues al adentrarnos en esta discusión laberíntica se corre el riesgo de perder de vista la cuestión. Esto es así porque una peculiaridad del fenómeno de la *akrasia* es su inquietante proximidad con otros fenómenos, con los que muchas veces puede confundirse o solaparse. A ello se suma una dificultad ulterior: pareciera que antes de definir satisfactoriamente la *akrasia* es preciso contar con teorías adecuadas sobre la naturaleza de los juicios morales, la agencia y la motivación, de la racionalidad en general y del razonamiento práctico en particular, entre otras cuestiones.

Las dificultades señaladas nos enfrentan al siguiente interrogante: ¿cómo tratar en forma conveniente un asunto tan complejo? Creo que ninguna estrategia podría resultar aquí *enteramente* satisfactoria. La complejidad del tema aunada al espacio limitado con el que contamos para abordarlo, nos impedirá de seguro ahondar en importantes tópicos conexos a la *akrasia*. Debemos contentarnos, en consecuencia, con un abordaje más modesto, aunque no por ello exento de rigurosidad. La naturaleza a priori imperfecta de este escrito no será obstáculo para emprender una tarea que aspira ante todo a la simpleza y la perspicuidad. A estos fines, se comenzará con una reconstrucción conceptual del problema (apartado 2) sin renunciar a sus aristas históricas (apartado 3), identificando así dos grandes enfoques en liza (3.1/3.2). En la sección subsiguiente (4), evaluaremos qué tipo de concepción de la *akrasia* puede mostrarse en especial fructífera. Finalmente, juzgaremos de qué mane-

---

1 Para un estudio de la *akrasia* en perspectiva histórica, véase Gosling (2003:7–96) y Hoffmann (2008).

ra, al enmarcarse en el dominio de la ética de la virtud, es posible arrojar luz sobre la *akrasia* desde una perspectiva más amplia (apartado 5).

### ¿Qué es la *akrasia*?

Comenzaremos delimitando nuestro objeto al mejor estilo socrático, esto es, intentando identificar aquellos aspectos esenciales a la definición. El término griego *akrasia* significa «falta de control» o «incontinencia», y se aplica típicamente a sujetos que pierden el dominio de sí mismos bajo ciertas circunstancias, por lo que suele vincularse también al fenómeno de la voluntad débil. Con base en esta razón, *akrasia* y debilidad de la voluntad han llegado a considerarse términos intercambiables. Aquí mantendré este criterio, sin desconocer eventuales matices (tal como veremos más adelante). Ahora bien, ¿qué es la *akrasia*? Como primera aproximación, puede decirse que ella hace referencia a la situación paradójica en la que un agente decide, libre e intencionalmente, actuar en contra de su mejor juicio. El carácter paradójico reside en que se opta por el peor curso de acción tras haberse decantado —a través de un proceso deliberativo— no por ese, sino por el curso de acción alternativo. Precisamente este aire de paradoja es lo que hace de la *akrasia* un predicamento tan difícil de explicar y sobre el cual los filósofos raramente concuerdan. Tales desacuerdos se suscitan no solo en lo que atañe a su definición, sino inclusive también en cuanto a la posibilidad de su existencia. Antes de adentrarnos en estos desarrollos, empero, procuremos delinear los elementos centrales que configuran la *akrasia*.

Según la definición mínima que se ha ofrecido, un sujeto actúa en forma acrática cuando ejecuta una acción determinada (y) *a sabiendas* de que existe una alternativa mejor (x). Aquí el acento recae en el conocimiento que el sujeto tiene acerca de la superioridad de un curso de acción (x) que, no obstante, termina por no elegir, privilegiando a su contrincante (y). La tradición filosófica ha tendido a interpretar este problema como un fallo en la racionalidad del agente. ¿De qué otra manera entender el que termine obrando en contra de su mejor juicio? Apelar a su irracionalidad se ha convertido así en moneda corriente a la hora de explicar el fenómeno. A ello acostumbra añadirse otra pieza explicativa crucial: el yerro de la razón se debe a que resulta desplazada o acallada por la pasión. No es casual que buena parte de quienes se abocaron a dilucidar el problema, lo hayan presentado como un conflicto entre la razón y las pasiones —o, en términos más acotados, entre el deber y el deseo—: la razón señala el derrotero a seguir, mas la pasión se interpone y consigue desviarnos. Si bien estos términos podrían matizarse en varios senti-



dos —como varios autores contemporáneos lo hacen—, resultará útil retener el esquema trazado a los presentes propósitos.

Es preciso establecer ahora bajo qué condiciones este esquema se torna operativo. Es decir, para que actuar en contra del mejor juicio constituya un caso de *akrasia* deben darse ciertas condiciones. En primer lugar, que se actúe libremente y no compelido por factores externos o internos. Ejemplo del primer tipo sería el de una acción inducida bajo amenaza, donde un sujeto debe realizar algo que considera incorrecto o contrario a su mejor juicio con un arma en su cabeza. Ejemplo del segundo tipo podría ser el de alguien severamente constreñido por factores psicológicos, tal como un adicto a las drogas demuestra ser cuando reconoce, por un lado, que sería mejor abstenerse de su consumo, siendo incapaz, por el otro, de hacerlo.<sup>2</sup> En segundo lugar, es preciso advertir que «ir en contra del mejor juicio» no hace referencia a lo que terceros estiman mejor hacer, sino a la perspectiva de la primera persona. Esto significa que el juicio en cuestión es realizado desde el punto de vista del propio agente, desde sus propias valoraciones y creencias (Mele, 2010:2). Finalmente, que los casos de *akrasia* no se reducen exclusivamente al dominio ético, esto es, a los juicios sobre lo mejor desde el punto de vista moral. En efecto, ya desde la antigüedad se ha tendido a teorizar sobre el fenómeno desde un marco mucho más amplio. Nadie diría, de esta suerte, que alguien incapaz de dejar de fumar es inmoral, aun cuando su conducta pueda catalogarse de acrática.

Recapitulando sobre lo dicho, es posible afirmar que un agente A actúa acráticamente cuando: a) primeramente considera, tras sopesar las razones relevantes, que es mejor hacer *x* que *y*; b) a la hora de actuar, A es libre de realizar tanto *x* como *y*; c) finalmente, A elige hacer *y* antes que *x*. Veamos todo esto con un ejemplo al que volveremos una y otra vez a lo largo de este trabajo. Supongamos que hemos llegado a convencernos de que sería mejor abstenernos de comer carne.<sup>3</sup> Imaginemos también que varias razones contribuye-

---

2 Si bien este ejemplo puede resultar controversial, es plausible afirmar que alguien que sufre de una seria adicción extendida en el tiempo, termina por perder en buena proporción su agencia y, por eso, su libertad. Una vez que el consumo de drogas se toma una práctica habitual, un estilo de vida incluso, las capacidades deliberativas y de agencia del sujeto se van desdibujando, en la medida en que se apropia de él un deseo compulsivo, cuasi mecánico. Las acciones compulsivas, luego, no serían equiparables a las acciones acráticas, ya que en estas últimas no es un deseo *irrefrenable* lo que parece estar en juego.

3 Este mismo ejemplo aparece en Holton (1999, 2003), y es comentado más tarde por otros autores (vg. Santamaría Navarro, 2008 y Mele, 2010—). En honor a la verdad, debo decir que había pensado en este evento puntual antes de leer el texto de Holton, razón por la cual decidí conservarlo a pesar de que pueda resultar manido. Por lo demás, lo considero un ejemplo atractivo por dos motivos principales: en primer lugar, porque representa bastante bien la complejidad entranada por los casos de *akrasia*; en segundo lugar, porque me interpela personalmente, al tratarse

ron a la formación de este juicio: razones éticas concernientes al status moral de los animales, razones ecologistas relacionadas con la contaminación producida por la cría industrial, razones prudenciales vinculadas con la propia salud y la conveniencia de llevar una dieta vegetariana. Es lógico que a partir de estas razones, nos planteemos: «sería mejor adoptar una dieta vegetariana y abandonar la dieta omnívora» (esto es, un juicio de la forma «es mejor hacer *x* que *y*»). Más tarde, con todo, flaqueamos ante la tentación de comer pollo al horno que nuestra madre ha preparado. Esta tesitura es acrática en la medida en que reúne los requisitos antes enumerados: se trata de una acción *no compélida* que contradice *nuestro* juicio acerca de lo mejor, y en donde lo mejor no hace referencia exclusiva a consideraciones morales. Adviértase también que el ejemplo propuesto entraña un conflicto entre la razón y el deseo.

Ahora bien, ¿cómo explicar la conducta paradójica y contradictoria que los casos de *akrasia* parecen envolver? A continuación, veremos que una de las grandes estrategias empleadas para resolver esta presunta paradoja ha consistido en la negación del fenómeno.

## **La *akrasia* en la tradición filosófica: dos enfoques contrapuestos**

La *akrasia* como fenómeno ilusorio

Para una venerable vertiente de la tradición filosófica, el problema entrañado por la *akrasia* es más aparente que real. De manera que la perplejidad que conlleva puede disolverse mediante la elucidación conceptual. Esta estrategia de análisis es expuesta primeramente por Platón en un extenso pasaje de *Protágoras* (1985:351–358).<sup>4</sup> Allí se busca clarificar —por boca de Sócrates— qué significa exactamente apartarse de lo que se considera mejor a causa del placer o de cualquier otra pasión<sup>5</sup> (1985:352 d–e). Seguidamente, se considera absurda semejante tesitura (1985:355 a–b) con base en una argumentación que aquí solo reproduciremos esquemáticamente. Sócrates parte de las siguientes tesis: a) todas las personas buscan su propio bien y rehúyen el mal; b) el bien se identifica con el placer y el mal con el dolor; c) los individuos actúan siem-

---

de una condición que experimenté tal cual se recoge en el ejemplo. Por fortuna, he conseguido vencer esos lapsos acráticos.

4 Un análisis exhaustivo sobre el tratamiento platónico de la *akrasia* en *Protágoras*, puede encontrarse en Gosling (2003:7–24).

5 En la antigüedad se pensaba que sobre todo los placeres vinculados con la comida, la bebida y el sexo tendían a desviarnos de lo mejor. Sócrates alude también a casos en que, a raíz de otras pasiones, alguien podría desviarse de ello, como por ejemplo un soldado que deserta del entrenamiento militar o un enfermo de su tratamiento a causa del dolor que le provocan (*Pr.* 354 a).

pre basados en lo que consideran mejor, esto es, en lo que *creen* les reportará mayor placer. De estas tesis se infiere que un individuo no puede apartarse de lo que considera mejor a causa del placer (ya que se ha asumido previamente que buscamos siempre el propio bien y que este se identifica con el placer). Lo que parece quedar en pie es más bien la siguiente posibilidad: que, enfrentado a la alternativa de elegir entre un placer mayor y un placer menor, termine por elegir el de menor realce. ¿Cómo explicar este fenómeno? ¿Acaso no termina de ser un caso de *akrasia*, en donde se escoge el peor curso de acción? Dadas las premisas de las que parte, Sócrates concluye que nadie puede escoger *voluntariamente* lo que considera peor:

nadie que sepa y que crea que hay otras cosas mejores que las que hace, y posibles, va a realizar luego esas, si puede hacer las mejores. Y el dejarse someter a tal cosa no es más que ignorancia, y el superarlo, nada más que sabiduría. (1985:358 c)

Según la tesis socrática, un individuo actúa *siempre* conforme a lo que considera mejor; solo que a veces, aquello que considera mejor, *realmente* no lo es. Luego, los casos en que se aparta de ello no se explican por su voluntad débil, sino por su ignorancia. Dada la concepción hedonista asumida, dicha ignorancia se vincula con un cálculo mal realizado en donde el agente sopesa incorrectamente placeres y dolores,<sup>6</sup> o placeres menores con placeres mayores.<sup>7</sup> Sócrates corona su razonamiento diciendo: «hacia los males nadie se dirige por su voluntad... ni cabe en la naturaleza humana, según parece, disponer a ir hacia lo que cree ser males, en lugar de ir hacia los bienes» (1985:358 d). El individuo supuestamente acrático no es tal. Su condición se parece más bien a la del autoengaño en la medida en que sustenta creencias ilusorias o erróneas. Si bajo esta luz consideramos el ejemplo planteado, esto es, el de quien vuelve a comer carne tras haberse propuesto no hacerlo, se trataría de alguien que cede ante ello porque cree (erróneamente) que *ese* curso de acción le reportará mayor placer —ignorando que, en el largo plazo, es la abstención cárnica la que le brindará mayor bienestar (una salud inquebrantable, una conciencia moral satisfecha, etc.).

En la filosofía contemporánea, Richard Hare sostendrá también la imposibilidad de la *akrasia*, enmarcando su argumentación en el prescriptivismo. Esta teoría envuelve una visión especial sobre la naturaleza de los juicios eva-

---

6 El caso del soldado que abandona su entrenamiento por pereza, o por temor (de ir a la guerra) exhibe un cálculo defectuoso: cree que la acción presente —desertar— le proporcionará mayor placer que no hacerlo. Semejante cálculo es, desde la perspectiva socrática, incorrecto, ya que a largo plazo esta acción terminará por causarle mayor pesar —vergüenza, culpa, etc.— que placer.

7 Platón abandona en diálogos posteriores la tesis hedonista expuesta en *Protágoras*.

lucativos, cuya instancia paradigmática está constituida por los juicios morales. Su rasgo central consiste en que apuntan a guiar la conducta (Hare, 1965:70), ya que surgen ante el interrogante «¿qué debo hacer?». De esto se desprende, según el autor, una conexión especial entre el juicio y la acción, de tal manera que cuando un agente afirma que debe realizar *x* se está comprometiendo a realizarlo. En los casos que entrañan una voluntad débil, encontramos precisamente en juego un juicio evaluativo conforme al cual un sujeto juzga que es mejor hacer *x* que *y*, y en donde finalmente se aparta de *x* para hacer *y*. En opinión de Hare, esto es sencillamente imposible:

es una tautología decir que no podemos asentir sinceramente a un imperativo que dirigimos a nosotros mismos, y *al mismo tiempo* no llevarlo a cabo, si es ahora la ocasión de hacerlo y se encuentra en nuestro poder (físico y psicológico) el realizarlo. (1965:79 – la traducción me pertenece)

Con esto se quiere decir que si un agente sostiene que debe hacer algo y reconoce que es ahora el momento de hacerlo, y no lo hace, entonces no estaba asintiendo *sinceramente* al juicio. Ello habilita interpretar cierto tipo de acciones «acráticas» como casos de insinceridad. El pasaje citado llama asimismo la atención sobre el siguiente punto: si una incapacidad física o psicológica impide realizar un curso de acción previamente estipulado como el mejor, entonces no podemos juzgarlo en calidad de acrático. Hare da a entender que una confusión de este tipo es lo que lleva a considerar acráticas acciones que en verdad no lo son. Como ejemplo menciona a Medea, quien presa de una pasión irrefrenable por Jasón, termina por hacer lo peor, siendo todavía capaz de entrever lo mejor. En la misma dirección se sitúan las palabras de San Pablo en *Carta a los Romanos VII*, quien se reconoce desgarrado por la ley del espíritu y la del pecado, mostrándose impotente para conducirse según lo primero a raíz de lo segundo. En ambas ocurrencias —concluye Hare— «no se encuentra en el poder psicológico de Medea ni de San Pablo actuar sobre la base de los imperativos implicados en los juicios morales que hacen» (Hare, 1965).<sup>8</sup> El filósofo inglés analiza también otra instancia que puede confundirse con un caso de voluntad débil: cuando una persona sostiene que «es mejor hacer *x* que *y*» desde la perspectiva de un tercero —vg., la sociedad (1965:83)—. Tales juicios son descriptivos antes que prescriptivos, y por ello no comprometen a la acción. Si alguien, al decir «es mejor hacer *x* que *y*» lo hace desde tal posicionamiento, su conducta no puede ser calificada de acrática al emprender después la acción contraria.

---

<sup>8</sup> En esta misma categoría ingresaría el caso del adicto a las drogas previamente aludido.

En opinión de Hare, los eventos por él analizados abarcan lo que suele mentarse mediante el concepto de «voluntad débil». Sin embargo, analizados de cerca, terminan mostrándose como otra cosa antes que como instancias de debilidad de la voluntad. Para seguir con nuestro ejemplo paradigmático: si un agente sostiene que debería dejar de comer carne y luego lo hace, entonces, desde la óptica de Hare, o bien no lo dijo sinceramente, o bien no lo dijo en primera persona —«yo creo que debería...»—, sino desde el punto de vista de algún otro —por ejemplo, la opinión de un defensor de los derechos animales cuyos libros recientemente leyó—. Al igual que Sócrates, Hare emplea el método del análisis filosófico para evidenciar que el desafío presupuesto por la cuestión de la *akrasia* se trata de un pseudoproblema.<sup>9</sup> Una vez desvelado su carácter ilusorio, lo que se ve emerger es un fenómeno diferente: ignorancia, falta de sinceridad, impotencia, o un uso equívoco del lenguaje.

### La posibilidad de la *akrasia*

En contraste con las voces escépticas que acabamos de sondear, otra importante vertiente admite la posibilidad de la *akrasia* y la integra al elenco de problemas filosóficos que con gran ahínco procurará diligenciar. Un examen cuidadoso se lo debemos a Aristóteles, quien fue el primero en subrayar la verosimilitud del fenómeno. En *Ética Nicomáquea* (*EN*) sostiene que la tesis de Sócrates se encuentra en manifiesta oposición con los hechos (*EN* 1145b:25–30), sentando así posición con respecto a la posibilidad de obrar a sabiendas contra lo mejor. Para echar luz sobre la cuestión, el estagirita se centra en la figura del *akrates* (o incontinente, tal como suele traducirse), de modo que su análisis hace foco en el carácter del agente. Para Aristóteles era fundamental, en orden al buen vivir, que un individuo desarrolle un buen carácter, un carácter virtuoso —como dirá—. Esta tarea, empero, no es nada fácil: requiere, entre otras cosas, de entrenamiento y habituación. Dado que nadie nace virtuoso, solo puede llegar a serlo con el tiempo a través de una dedicación esmerada y constante. Esto explica que muchos individuos se encuentren a mitad de camino, por así decir; todavía no han alcanzado la virtud completa, y este parece ser el caso del *akrates*. El mismo, en efecto, está bien

---

<sup>9</sup> Por lo demás, esta conclusión se sigue lógicamente de la estrechísima conexión que ambos establecen entre juicio y acción, esto es, de la capacidad que el conocimiento tiene de motivar *inmediata* y *directamente* a la voluntad. Esta postura suele catalogarse como «internalismo extremo». Para un análisis de posturas internalistas y externalistas respecto de la cuestión de la *akrasia*, véase Stroud y Svirsky (2019). Para un texto clásico que explica la distinción entre razones externas e internas, véase Williams (1993).

encaminado: sabe hacia dónde ha de dirigirse, solo que a veces se distrae y equivoca el camino. Aristóteles emplea una metáfora vívida: el incontinente se parece a una ciudad que, a pesar de decretar buenas leyes, no las aplica (EN 1152a:20). Con ello se sugiere que aquel posee los principios generales o las creencias apropiadas con respecto a la acción, pero yerra en su ejecución. Precisamente porque la pasión logra subyugarlo.

Aristóteles presenta a la *akrasia* como una circunstancia en donde el sujeto se halla tironeado por la razón y los apetitos. El incontinente sabe que obra mal por la pasión (EN 1145b:10): posee las actitudes teóricas correctas hacia el placer, pero es débil frente al mismo (Rorty, 1980:275). Nuestro filósofo se propone examinar hasta qué punto el agente en cuestión obra con conocimiento al ceder a la pasión, y en esto parece seguir la pista dejada por Sócrates.<sup>10</sup> Aquellos dominados por la pasión poseen el conocimiento de lo mejor, pero al modo de un principiante en ciencias que ensarta frases sin saber lo que dice, «pues hay que asimilarlo y esto requiere tiempo» (EN 1147a:20). Esta afirmación se halla perfectamente a tono con el desarrollo lento que requiere la virtud: el conocimiento de lo bueno no es suficiente por sí solo para motivar la conducta; para ello es preciso su persistente puesta en práctica hasta lograr una completa interiorización. Al claudicar ante la pasión, el conocimiento del *akrates* se muestra vacilante e imperfecto al carecer de integración con su praxis.

¿No volvemos así a la tesis socrática? Es decir, a la tesis con arreglo a la cual no es una voluntad débil lo que nos aparta de lo mejor, sino la ignorancia. Hay dos elementos que nos permiten dudar fundadamente de ello. En primer lugar, que Aristóteles complejiza la visión de la *psyche* humana desafiando el intelectualismo extremo de Sócrates. Para este, no existía hiato alguno entre conocimiento y voluntad: basta tener una creencia sobre cómo alcanzar algo que se considera bueno —un juicio evaluativo, al decir de Hare— para poseer la motivación correspondiente. De allí que no pueda existir tal cosa como una voluntad débil, porque esta *siempre* se pliega al conocimiento (a la creencia que el sujeto posee). Aristóteles, en cambio, piensa que la voluntad o, mejor dicho, los deseos, son capaces de desoír la razón cuando se encuentran ineducados —tal como se evidencia en el caso del *akrates*—. La integración completa entre ambos es posible, pero solo a fuerza de práctica y constancia. En segundo lugar, y este aspecto lo distancia más decisivamente de Sócrates, Aristóteles dispone de un marco conceptual que permite responsabilizar al *akrates* por su conducta, ya que se trataría, en definitivas cuentas, de un tipo de acción *voluntaria*. Dado que el mismo *sabe* en qué clase de situación es capaz de perder el dominio de sí y rendirse a la pasión, no debería, entonces, exponerse a

---

10 Es decir, se pregunta si es que acaso no se actúa así por ignorancia.

tal coyuntura. Si lo hace, se vuelve responsable de ello al colocarse en un trance que *voluntariamente* podría haber evitado (Rorty, 1980:279).

La posición aristotélica es, claro está, más sutil y compleja de lo que hemos mostrado. Pero baste con lo dicho a los presentes propósitos. Veamos ahora cómo nuestro ejemplo por antonomasia sería susceptible de análisis bajo la óptica de Aristóteles: si sé de antemano que el deseo de comer carne continúa asediándome, a pesar del juicio que me he formado sobre la conveniencia de abandonarla, entonces debería evitar aquel tipo de situaciones en donde ese deseo podría activarse —excusándome, por ejemplo, de asistir a una reunión donde sé que se servirá ternera—. Si no lo hago y me rindo luego a la tentación de degustarla, mi comportamiento es genuinamente acrático y soy responsable del mismo.

En los confines de la filosofía contemporánea se multiplicaron las tentativas tendientes a explicar el fenómeno de la *akrasia*. Por el impacto que provocaron, aquí recogeremos las interpretaciones que Donald Davidson y Richard Holton han ofrecido. Gracias a un sugerente artículo publicado en 1970, «¿Cómo es posible la debilidad de la voluntad?», el primero consiguió revitalizar las discusiones sobre el tópico. Allí declara que a pesar de la resistencia de Hare en reconocer la existencia de acciones acráticas, a él le parece absolutamente cierto que existen (Davidson, 1995:46). Igualmente considera que las explicaciones ensayadas por la tradición son insatisfactorias, por lo que avanzará una lectura original del tema. En su opinión, el problema debe enmarcarse en el dominio de la filosofía de la acción y ha de partir de una concepción adecuada del razonamiento práctico. A este respecto despliega la siguiente definición:

Al hacer *x*, un agente actúa de manera incontinente si y sólo si: (a) el agente hace *x* intencionalmente; (b) el agente cree que hay una acción alternativa y que le es asequible; y (c) el agente juzga que, hechas todas las consideraciones, sería mejor hacer *y* que hacer *x*. (Davidson, 1995:38)

El elemento novedoso yace en la expresión «hechas todas las consideraciones», y se convertirá en pilar para que Davidson erija su interpretación. En una fase temprana del razonamiento práctico, el agente suele evaluar los pros y contras que atañen a los posibles cursos de acción. Así, en situaciones de conflicto en donde existen dos alternativas en liza, como típicamente lo encarnan los casos de incontinencia, un sujeto evalúa qué razones soportan cada uno de ellos. Tras sopesarse las razones de un lado y de otro, se obtiene el juicio «sería mejor hacer *y* que *x*», el cual, en opinión de Davidson, se trata de un juicio relacional en donde las razones que sustentan hacer *y* antes que *x* descansa en su

superioridad con respecto a un conjunto definido de consideraciones, esto es, aquellos aspectos estimados relevantes. Esto no significa, luego, que *ninguna razón* apoye a la alternativa considerada de menor calibre. Con esta propuesta, Davidson desafía la imagen tradicional de la *akrasia* como un conflicto entre la razón y la pasión, poniendo especialmente en cuestión la concepción aristotélica (Davidson, 1995:50–54). La idea es que la oposición se da siempre entre un conjunto de razones y otro conjunto de razones:

Es bastante claro que la incontinencia puede existir sólo cuando hay conflicto en este sentido, pues el hombre incontinente sostiene que un curso es el mejor (por una razón) y, sin embargo, hace alguna otra cosa (también por una razón). (Davidson, 1995:52)

Para ilustrar la posición davidsoniana a partir de nuestro ejemplo, diríamos que hay razones tanto para seguir una dieta vegetariana como para seguir una dieta omnívora. Entre las que apuntalan esta última opción, contaría el que es eventualmente más económica; el que es de más fácil ejecución, ya que una dieta vegetariana requiere bastante conocimiento sobre cómo combinar alimentos para evitar déficits nutricionales; y el que también es más conveniente desde un punto de vista social, evitándonos así sentir cierta incomodidad en reuniones familiares y de amigos, en las que suelen servirse platos que contienen carne. Seguramente estas son razones, pero no —tal como se ha planteado en el ejemplo— razones adecuadas o relevantes para dirimir la cuestión. Recordemos que su relevancia fue establecida desde un punto de vista ecológico, nutricional y moral. Si hechas todas las consideraciones, concluimos que es preferible la dieta vegetariana, pero luego nos inclinamos por comer carne, actuamos de manera incontinente según Davidson: pues fallamos en formarnos la motivación correspondiente al mejor juicio.<sup>11</sup> Todavía más: actuamos en forma irracional puesto que si bien hay razones para optar por una dieta omnívora, estas razones no serían *suficientes* —según las consideraciones que nosotros mismos hemos hecho (Stroud y Svirsky, 2019).

En función de lo expuesto hasta aquí, tanto Aristóteles como Davidson coinciden en ver la *akrasia* como cierto fallo racional, solo que explicado en direcciones diferentes: en el caso de Aristóteles, ese fallo se explica basado en la pasión que arrebatada y nubla el juicio, mientras que Davidson lo enfoca como una instancia en donde el mejor juicio no consigue fuerza motivacional. Richard Holton merece mención aparte al plantear una interpretación

---

11 Davidson llama «principio de continencia» al principio sustantivo que oficiaría de enlace entre el mejor juicio —todo considerado— y la acción. El incontinente carece de este principio (Davidson, 1995:61).



del fenómeno del todo novedosa, alcanzando una influencia destacable. En un artículo publicado a fines de los 90, el filósofo norteamericano propone tratar a los términos *akrasia* y «debilidad de la voluntad» como términos independientes. En su opinión, la tradición los ha empleado en forma intercambiable porque ha tendido a confundir dos condiciones diferenciables:

Los casos centrales de debilidad de la voluntad se caracterizan mejor, no como casos en que la gente actúa en contra de su mejor juicio, sino como casos en donde fallan en actuar con arreglo a sus intenciones (...) Voy a sugerir que los casos de debilidad de la voluntad surgen cuando los agentes tienden con bastante facilidad a reconsiderar sus intenciones. (Holton, 1999:1 – la traducción me pertenece)

Como puede verse, Holton no descarta ni niega el hecho de la *akrasia*. Lo que propone, más bien, es diferenciarlo del estado de la voluntad débil, cuyo rasgo neurálgico se aprecia en especial si se concibe, no tanto como el desvío del mejor juicio, sino como el abandono de una *intención* previamente formada. Con ello se introduce un elemento hasta entonces pasado por alto en los análisis clásicos del tema. En estos, en efecto, habían sido las creencias y los deseos los estados mentales principalmente considerados. Pero ahora Holton concederá a las intenciones —actitudes con contenido proposicional— el rol protagónico. Una de las ventajas de este enfoque consiste en su fertilidad explicativa en situaciones donde la explicación tradicional fracasa. Por ejemplo, los casos de conflictos entre normas o valores inconmensurables. En semejante tesitura un agente se enfrenta a una disyuntiva compuesta por alternativas incomparables; por consiguiente, la pretensión de identificar el *mejor*<sup>12</sup> curso de acción está de antemano destinada a naufragar. Holton alude al caso narrado por Sartre en *El existencialismo es un humanismo*, en el que un individuo debe elegir entre pelear contra el fascismo o quedarse en su casa para cuidar de su madre. En esta situación, no parece haber manera de dirimir racionalmente la cuestión, esto es, de formarse un juicio justificado respecto del mejor curso de acción; motivo por el cual solo cabría una elección radical. O en términos de Holton, *formarse una intención*: supongamos que sea la de alistarse en el ejército y pelear contra el fascismo. Si luego el mismo agente decide quedarse en su casa, abandonando así su intención inicial, puede ser acusado de poseer una voluntad débil (Holton, 1999:8).

---

12 Puesto que «mejor» se trata de un término relacional o comparativo, en los casos de conflictos inconmensurables seríamos incapaces de determinar el mejor curso de acción, debido a que las alternativas en juego, o bien carecen de un patrón común de medida, o bien dificultan establecer un meta-criterio que permita ordenarlas en una secuencia de prioridades. A este respecto, véase Lariguet (2010).

Veamos paso por paso el argumento. Para empezar, el autor en cuestión define a la intención como a una pieza de carácter motivacional que impulsa a un agente racional a actuar. Es decir, las intenciones se forman sobre la base de propósitos que procuramos alcanzar a través de la acción. Es característico de las instancias en que se exhibe una voluntad débil, el abandono de una intención previamente asumida, la discontinuidad de un proyecto, o la falta de constancia en la persecución de una meta. Ahora bien, esto no significa que todos los eventos en que se abdica de una intención puedan ser catalogados de ejemplos de voluntad débil. Para ello deben darse ciertas condiciones, y en especial, son las siguientes: cuando se tiende a revisarla o abandonarla con demasiada facilidad y, sobre todo, cuando no hay una buena razón para hacerlo. Por ello Holton dará un paso más y dirá que los casos de voluntad débil se identifican por una *revisión irrazonable* de la intención. Esto significa, por oposición, que si un individuo se niega a reexaminar un propósito a la luz de un motivo justificado para hacerlo, no estamos ante un caso de fortaleza o constancia, sino de obstinación (Holton, 1999:8).

En una precisión ulterior de la definición, el autor añade que las intenciones son resoluciones cuya función, en parte, consiste en vencer las inclinaciones contrarias que uno teme podrían aparecer (Holton, 2003:42). Esto implica que al resolverse a hacer algo, un individuo está también comprometiéndose a repeler deseos o impulsos que podrían desviarlo de su meta. Si estas inclinaciones opuestas lo asedian, pero resiste, entonces nos encontramos ante un caso de fuerza de voluntad. En virtud de estas apreciaciones se concluye que: «la debilidad de la voluntad es, aproximadamente, la revisión irrazonable de una resolución; la fuerza de voluntad es su mantenimiento razonable» (Holton, 2003:61). A partir de estos desarrollos, Holton cree haber arrojado luz sobre aspectos soterrados en las discusiones sobre el fenómeno de la voluntad débil, al tiempo que auspicia su diferenciación de la *akrasia*. Este término quedaría reservado para los casos en que una acción va a contramano del mejor juicio. Al separar ambas condiciones, una acción podría considerarse acrática sin evidenciar una voluntad débil, y viceversa. Así, por ejemplo, si una persona llega a convencerse que debería dejar de comer carne sobre la base de buenas razones, pero sin formarse la intención respectiva, no podríamos acusarla de poseer una voluntad débil en caso de no hacerlo, aunque sí de actuar acráticamente (Holton, 2003: 41). A la inversa, alguien podría considerar que hay mejores razones para adoptar una dieta omnívora, resolviéndose no obstante ello a implementar una dieta vegetariana. Si luego abandona injustificadamente esta última, su conducta evidenciaría una voluntad débil sin ser acrática.

## La *akrasia* como concepto polifacético

Si hay algo que puede colegirse en función de los contenidos precedentes, es que las disquisiciones filosóficas sobre la *akrasia* distan de ser concordantes. Adicionalmente, la literatura que recoge este debate es abundante, de manera que el recorrido que hasta aquí se ha trazado es apenas una muestra ínfima de ese universo vasto y complejo. Con todo, es preciso adoptar —aunque sea provisionalmente— algún tipo de definición. Es así que en este apartado propondré, a título exploratorio, que el concepto de *akrasia* se trata de un concepto polifacético, capaz de dar cuenta de un conjunto variado —y potencialmente abierto— de casos. Claro está que, no obstante su variedad, estos casos poseen un «parecido de familia» sin el cual difícilmente agruparíamos bajo un mismo concepto. Con este objetivo en miras, se evaluarán los aportes de los distintos autores considerados hasta el momento, así como los de otros que oportunamente se introducirán.

La primera cuestión a dirimir es la de si los términos *akrasia* o debilidad de la voluntad hacen referencia a un fenómeno real o si, por el contrario, describen una situación imposible. Después de todo, podrían tratarse de conceptos que ofrecen una caracterización errónea de los eventos que se buscan examinar (Santamaría Navarro, 2008:221-222). Hemos visto sumariamente los argumentos de Sócrates y Hare para negar la realidad del fenómeno, y si bien sus análisis son sutiles y firmes desde un punto de vista lógico, sus explicaciones se limitan a un número restringido de casos que no agotan su complejidad.<sup>13</sup> Por lo demás, desde un punto de vista fenomenológico, parece indudable que tales hechos existen. Si cada uno dirige la mirada a su propia interioridad, ¿cuántas veces nos pasó que terminamos haciendo algo que consideramos mejor no hacer, o viceversa?; ¿quién de nosotros no se propuso alguna vez comenzar a ejercitarse en un gimnasio, para posponer el mismo propósito una y otra vez?; ¿cuántas veces nos excedimos en la comida y la bebida, a pesar de haber considerado mejor hacerlo dentro de ciertos límites? No dudaríamos en denominar tal tipo de episodios «acráticos», o en considerar que corresponden a una voluntad débil.

Adviértase que a pesar de Holton, se ha decidido continuar empleando *akrasia* y «debilidad de la voluntad» como términos equivalentes. Diferenciarlos, en mi opinión, complica innecesariamente el paisaje conceptual,

---

13 Sócrates restringe la discusión a casos donde un sujeto es supuestamente arrastrado por la pasión; Hare, por su parte, principalmente a casos en donde aquel es inerte frente a fuerzas indomables (las propias inclinaciones). Asimismo, la aceptación de sus conclusiones requiere de la aceptación previa de una tesis internalista fuerte con respecto a los juicios evaluativos.

dado que, en definitiva, no parecen aludir a fenómenos *completamente* independientes. Es decir, normalmente un agente se forma la intención de hacer  $x$  porque ha juzgado que es mejor hacer  $x$ . Es *usual* que los juicios vayan acompañados de intención, y que las intenciones entrañen un juicio evaluativo — aunque podrían detectarse excepciones, tal como el caso mencionado de los conflictos inconmensurables—. El mismo Holton reconoce más tarde que la cuestión es más compleja de lo que había inicialmente pensado, cosa que le lleva a matizar su posición y a reconocer que las instancias de *akrasia* —en las que se actúa en contra del mejor juicio— deberían considerarse instancias de debilidad de la voluntad (Holton y May, 2012:342); con lo cual pierde fuerza la separación nítida que previamente se había postulado entre ambas. Con arreglo a estas consideraciones, se muestra más adecuada una estrategia que busca integrar en el mismo concepto, las dos dimensiones.

De este modo, los casos de *akrasia* o debilidad de la voluntad, podrían describirse como casos en que un agente  $A$  juzga mejor hacer  $x$  que hacer  $y$ , y termina pese a ello haciendo  $y$ ; o como casos en que  $A$  tiene la intención de hacer  $x$  y evitar hacer  $y$ , pero termina decidiéndose en favor de  $y$ ; o como casos en donde convergen ambas condiciones. Según las explicaciones de los diferentes autores examinados, este fallo puede deberse a que el agente resulta momentáneamente distraído por un deseo (Aristóteles), o a que es incapaz de actuar acorde a las mejores razones detectadas (Davidson), o a que no atina a perseverar en un propósito (Holton). Estas interpretaciones realizan valiosos aportes desde diferentes ópticas, y pueden ser incorporadas armónicamente en el seno del mismo concepto. Esto es así porque confluyen en la explicación de un fenómeno que más allá de las múltiples facetas que pueda asumir, evidencia siempre un rasgo en especial: la incapacidad de un sujeto de dominarse a sí mismo, la ineptitud para alcanzar una meta o continuar con un proyecto, no porque fuerzas externas e inmanejables se lo impidan, sino porque cierta falencia de la que él mismo es responsable se lo dificulta.

Se ha visto también que dicha falencia ha solido encuadrarse como un fallo racional. Sin entrar a examinar tan ardua cuestión, aquí se asumirá que la irracionalidad o la falta de racionalidad constituyen *parte* de la explicación del fenómeno, sin agotar su totalidad. Por eso son bienvenidos otros enfoques que sacan a la superficie ribetes adicionales. En este sentido, considero particularmente interesante una línea de análisis que estima a la *akrasia* no tanto como un problema asociado con la racionalidad, sino con las costumbres. Inspirándose en la filosofía de William James, Bromhall (2018) considera que los casos de *akrasia* se explican mejor cuando son enfocados como un tipo de acción que procura ir en contra de un hábito establecido, pero que falla en ese intento. En otras palabras: un individuo adquiere patrones de com-

portamiento del entorno o la sociedad en la que vive, configurando así hábitos y costumbres que le permiten resolver con celeridad diversas situaciones que se le presentan. Cuando este mismo individuo se forma un juicio que requiere actuar en contra de las costumbres dadas —y que él mismo ha interiorizado—, muy probablemente no lo consiga, dada la fuerza motivacional que aquellas han logrado. Pensemos cuán apropiada resulta también esta óptica para analizar nuestro ejemplo: el de alguien que falla en seguir una dieta vegetariana *porque durante mucho tiempo ha tenido el hábito de consumir carne*.

El enfoque centrado en la fuerza de las costumbres arroja sin dudas luz sobre el fenómeno. No obstante, sería un error pensar que todos los casos de voluntad débil pueden reducirse a dicha explicación.<sup>14</sup> Una vez más, la adopción de una táctica integral que evite apreciaciones unilaterales, continúa mostrándose mayormente apropiada. Por último, un aspecto adicional merece atención: la *akrasia* ha tendido a enfocarse como un estado en el que dos actitudes mentales —creencias y deseos— entran en conflicto *en el mismo momento*. Según Monika Betzler, dicha descripción no satura todas las posibilidades:

Me gustaría, en contraste con la concepción de estados sincrónicos, introducir la concepción en términos de procesos diacrónicos de la debilidad de la voluntad. Dicha concepción concierne a casos en los que, tras una ponderación de razones, nos comprometemos con una determinada alternativa de acción. Llegamos a una conclusión de este proceso, y lo hacemos formando una intención de perseguir dicha acción o un fin determinado. En el transcurso del tiempo se demuestra, sin embargo, que no podemos mantenernos en ello. Actuamos con voluntad débil en la medida en que renunciamos a este compromiso... (2009:197)

A continuación, se verá por qué esta concepción que atiende a los procesos diacrónicos es relevante. Ella nos permite insertar el fenómeno de la voluntad débil en un encuadre más amplio: el de la formación del carácter como un proceso que abarca buena parte del transcurso vital.

---

14 Por ejemplo, alguien que nunca fue infiel, podría actuar acráticamente tras considerar que sería mejor no engañar a su pareja, y luego hacerlo. Aquí el factor explicativo no yace en sus hábitos, al tratarse de una persona que *no acostumbraba* ser infiel.

## Una aproximación desde la ética de la virtud

A modo de cierre, me gustaría mostrar cómo nuestra visión sobre la *akrasia* puede enriquecerse si se enfoca desde una orientación particular: la ética de la virtud. Con ello no asumo que esta es la única o la mejor forma de enfocarla, sino más bien que al hacerlo ganamos una comprensión más amplia del problema. Las ventajas de apreciar el fenómeno de la *akrasia* bajo las lentes de la ética de la virtud, consiste en que permite integrar varios de los elementos analizados y evidenciar por qué, en definitiva, es importante que un individuo pueda dominarse a sí mismo. En la escena contemporánea, han predominado los abordajes que tratan la debilidad de la voluntad como un problema concerniente a la filosofía de la acción; sin embargo, esto no ha sido siempre así: tanto en la antigüedad como en el medioevo, se divisaba como una dificultad asociada al carácter del agente. Y en ello hacen foco, precisamente, las éticas de la virtud: en el desarrollo de un buen carácter o de una manera de ser virtuosa, de la que Aristóteles es gran exponente.

A diferencia de las teorías modernas, que tienden a concebir la moralidad en forma más restrictiva,<sup>15</sup> en las éticas de la virtud se plantean un conjunto de preocupaciones más amplias, en donde no solo importa cómo vincularse con rectitud frente a los otros, sino igualmente con respecto a uno mismo. El agente capaz de compaginar con destreza ambas dimensiones, es aquel que alcanza una vida bien lograda.<sup>16</sup> Desde esta singular perspectiva, la vida buena no se logra por azar, sino a través de un trabajo denodado sobre sí mismo: forjando un cierto tipo de carácter a través de la práctica de las virtudes. Estas, según Aristóteles, corresponden a excelencias intelectuales y del carácter que nos permiten pensar, deliberar y actuar en forma señalada. Su valor es doble, ya que además de poseer un valor intrínseco, nos permiten alcanzar fines valiosos. La valentía, por ejemplo, es un rasgo del carácter en sí admirable, pero también una cualidad que nos permite alcanzar fines estimables: salir airosos frente al peligro, o soportar con entereza las contrariedades que eventualmente se presentan. Lo mismo cabe decir con respecto al resto de las virtudes: justicia, generosidad, templanza, prudencia, etcétera.

Según esta caracterización, un individuo raramente alcanzará las diversas metas que a lo largo de su vida se propone, si no desarrolla las virtudes. Aquí

---

15 Como la esfera de las obligaciones morales, esto es, de aquello que nos debemos los unos a los otros.

16 Los griegos denominaban a semejante condición *eudaimonia*, de la que «felicidad» es una traducción insuficiente. «Vida buena» o «florecimiento humano» constituyen, quizá, expresiones más afortunadas.

no necesitamos comprometernos con tal tesitura,<sup>17</sup> aunque sí podemos afirmar que la consecución de cualquier proyecto vital requiere, para su ejecución exitosa, de autodominio. Diversos estudios en psicología demuestran su importancia para el logro de metas de mediano y largo aliento (Landinez Guio, 2019:2-4).<sup>18</sup> Si evocamos a este respecto la figura del *akrates*, podemos ver en qué reside exactamente su defecto: en el escollo que suele encontrar para llevar a buen puerto sus propósitos. Su situación no es, con todo, insuperable. Con arreglo a la ética de la virtud de corte aristotélico, las instancias en que actúa acráticamente son ocasión para la reflexión y la modificación de la conducta —lo cual revertirá, a la postre, en la modificación de su carácter—. Aristóteles concedía que el *akrates* no está lejos de la virtud: al poseer el conocimiento de lo mejor, se da cuenta cuándo yerra y por qué yerra, cosa que le conduce al arrepentimiento (*EN* 1151a:10-15). De esta suerte, y por más extraño que suene, la *akrasia* poseería un valor positivo al convertirse en acicate para el progreso personal, en cuanto mejora del propio carácter. Dado que el buen carácter se forja a través del tiempo —a través de procesos diacrónicos, a decir de Betzler—, las ocasiones en que fallamos nos permitirán realizar los ajustes pertinentes. Paradójicamente, la *akrasia* es un desperfecto que contiene las condiciones para su propia superación.

---

17 Es decir, con la idea de que necesitamos desarrollar *todas* las virtudes para alcanzar una vida bien lograda. Aristóteles creía, además, en la tesis de la unidad de las virtudes.

18 Apoyándose también en diversos estudios de esta naturaleza, Holton sostiene que el autodominio es un buen predictor de éxitos futuros relacionados con excelencias sociales y académicas de amplio espectro (Holton, 2003:54).

## Referencias bibliográficas

- Betzler, M. (2009). Debilidad de la voluntad como irracionalidad furtiva. *Ideas y Valores*, (141), 191–215.
- Bromhall, K. (2018). Embodied Akrasia: James on motivation and weakness of will. *William James Studies*, 14(1), 26–53.
- Davidson, D. (1995). ¿Cómo es posible la debilidad de la voluntad? *Ensayos sobre acciones y sucesos* (pp. 37–62). UNAM.
- Gosling, J. (2003). *Weakness of the Will*. Routledge.
- Hare, R. (1965). *Freedom and Reason*. Clarendon Press.
- Hoffman, T. (Ed.) (2008). *Weakness of Will from Plato to the Present*. The Catholic University of America Press.
- Holton, R. (1999). Intention and Weakness of Will. *The Journal of Philosophy*, 96(5), 241–262.
- Holton, R. (2003). How is Strength of Will Possible? En Stroud, S. y Tappolet, C. (Eds.), *Weakness of Will and Practical Irrationality* (pp. 39–67). Oxford University Press.
- Holton, R. y May, J. (2012). What in the world is weakness of will? *Philosophical Studies*, 157(3), 341–360.
- Landinez Guio, D. (2019). Debilidad de la voluntad y autocontrol: una discusión en torno a la teoría de las funciones ejecutivas. *Límite. Revista Interdisciplinaria de Filosofía y Psicología*, 14(5), 1–11.
- Lariguet, G. (2010). Los dilemas morales *qua* límites de la racionalidad práctica. *Diánoia*, LV(64), 71–108.
- Mele, A. (2010). Weakness of Will and Akrasia. *Philosophical Studies*, 150(3), 391–404.
- Platón. (1985). *Protágoras*. Gredos.
- Rorty, A. (1980). Akrasia and Pleasure. En Rorty, A. (Ed.), *Essays on Aristotle's Ethics* (pp. 267–284). University of California Press.
- Santamaría Navarro, V. (2008). Akrasia vs. Debilidad de la voluntad. *Daimon. Revista de Filosofía*, Suplemento 2, 211–222.
- Stroud, S. y Svirsky, L. (2019). Weakness of Will. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2019/entries/weakness-will/>
- Williams, B. (1993). Razones Internas y Externas. En *La Fortuna Moral* (pp. 131–245). UNAM.



### **III. Recientes proyecciones sobre problemas persistentes**

## 19. Una lectura wittgensteiniana del campo de la Ética

María Sol Yuan

La paradoja desaparece sólo si rompemos radicalmente con la idea de que el lenguaje funciona siempre de un solo modo, sirve siempre para la misma finalidad: transmitir pensamientos —sean éstos luego sobre casas, dolores, lo bueno y lo malo, o lo que fuere.

*Wittgenstein*

¿Cómo podemos justificar un conjunto de juicios morales frente a otro diferente sobre el mismo asunto? ¿Cómo podemos elegir entre dos conjuntos de juicios morales distintos? Estos dilemas vinculados al desacuerdo epistémico-moral y a las elecciones racionales revisten un interés que excede el marco de la Ética y la disciplina filosófica en general. Quiero apresurarme a desilusionar al lector confirmando mi incompetencia para responder a estos dos interrogantes que me inquietan (recomiendo, en este sentido, la contribución de Martínez Zorrilla a este libro). A pesar de ello, creo que hay algunos elementos que se encuentran en el punto de partida de estos problemas y que bien valdría la pena considerar. Discutiré, en particular, dos representaciones de cómo puede concebirse el campo de la ética: como algo individual (centrado en la autenticidad, la libertad, la responsabilidad, la motivación, o lo que se quiera mencionar, pero vinculado siempre a un aspecto determinadamente individual de las personas) y como conteniendo reglas (más o menos) universales y a priori. Ejemplificaré brevemente qué tengo en mente cuando me refiero a estas dos ideas.

*Justice for Hedgehogs* (2011), de Ronald Dworkin, puede servir de ejemplo de una caracterización de la ética como algo individual. A pesar de que una de sus principales tesis es que la ética y la moral dependen la una de la otra, el campo de cada una es una región distinguible entre lo individual y lo comunitario. Según sus propias palabras, su intento es el de «ilustrar, al menos, la unidad de los valores éticos y morales: describo una teoría de cómo es vivir bien y, si queremos vivir bien, qué debemos hacer y no hacer por otras personas» (Dworkin, 2011:1). Tenemos, entonces, un conjunto con una definición de en qué consiste vivir bien como una meta individual a la que puede aspirar una persona y, por otro lado, las acciones a las que me obligo con relación

a los otros. Luego, si quiero vivir bien, debo hacer tal y tal cosa en relación con los otros.

La segunda idea a la que deseo oponerme consiste en la afirmación de que el campo de la ética puede ser delimitado por reglas universales y a priori. Como ejemplo, mencionaré la posición encarnada por Sócrates en *República* (Libro II) al hablar de la justicia en términos de un estándar independiente y eterno contra el cual las prácticas humanas deben ser juzgadas. Sócrates rechaza la propuesta de Glaucón de considerar a la moral como una creación humana, una regla que las personas introducen para garantizar la estabilidad pacífica de la sociedad.

Estos dos aspectos juegan un rol decisivo si uno quiere adentrarse en preguntas como las que mencioné al comienzo de esta introducción respecto al desacuerdo moral y diferencias culturales sobre propiedades morales como «bueno», «correcto», «incorrecto», «justo», etc. La razón es que aceptar o rechazar la dimensión comunitaria y la falta de un suelo firme a la hora de pensar en la atribución de propiedades morales impacta en la propuesta de alternativas para la solución de estos dilemas vinculados a la justificación de juicios evaluativos.

A continuación, intentaré defender, en primer lugar, la idea de que el campo de la ética puede ser comprendido de manera más compleja apelando a una lectura comunitarista de las reglas que la gobiernan. En segundo lugar, mencionaré que el consenso no elimina los hechos sino, solamente, su interpretación metafísica respecto del vínculo que las propiedades morales mantienen al ser empleados en enunciados indicativos. Por lo tanto, el campo de la ética puede verse integrado por un conjunto de normas prácticas comunitariamente establecidas, sobre hechos que no responden a esencias o realidades supra humanas sino al modo «amable» en que se comporta la realidad en nuestro vínculo contextualizado y siempre particular.

El esquema de desarrollo que seguiré será el siguiente: en el primer apartado presentaré cómo puede entenderse a las reglas que conforman el campo de la ética desde una perspectiva comunitarista, extendiendo el diagnóstico de Kripke de la solución a la paradoja del seguimiento de reglas de Wittgenstein, al campo de la ética en cuestión. En particular, haré uso del mencionado «juego de atribución de conceptos» explicitado por Kripke. En el segundo apartado de este trabajo, tomaré distancia de la posición de este autor que sostiene que la solución a la paradoja en el seguimiento de reglas ha de ser «escéptica», es decir, debe partir de la aceptación de que no hay ningún hecho que me permita determinar el significado de los conceptos. Aquí, intentaré mostrar que no hay ningún hecho «mental» privado en función del cual poder asegurar el empleo (y el significado) de los conceptos para una infinidad

de posibles aplicaciones, pero que es posible pensar en hechos que se centren en nuestras prácticas y formas de vidas, en el modo en que se vincula la realidad con nuestros empleos estables.

No puedo abandonar este apartado sin aclarar, brevemente, algunas cuestiones vinculadas a la elección del tema y a la perspectiva con la que se emprende su estudio. Esta aclaración podría resumirse diciendo que no realizaré ningún aporte a los debates sustantivos particulares que suelen incluirse en la Ética.<sup>1</sup> En cambio, consideraré desde una perspectiva metateórica el campo de la ética, es decir, qué es aquello que estamos señalando cuando afirmamos que tales y tales hechos caen bajo el dominio de la Ética y pueden, por ende, ser juzgados como moralmente correctos o incorrectos.

La perspectiva wittgensteiniana escogida para este estudio hace que lo que hasta ahora es una intromisión en esta subdisciplina filosófica se torne una cabal imprudencia, habida cuenta de que a pesar de lo que suele mencionarse la relevancia de la ética para el pensamiento de Wittgenstein, sus escritos filosóficos no suelen ahondar (por diferentes razones de su autor) sobre el tema y cuando lo hacen, suelen ser en un sentido bastante crítico respecto a la posibilidad de establecer juicios deliberativos en ética.<sup>2</sup>

A pesar de lo dicho, las dudas que presento se relacionan con los supuestos epistemológicos, metafísicos y semánticos que sirven de andamiaje para discusiones sustantivas en el campo de la Ética. Además, creo que las herramientas que provienen del segundo Wittgenstein pueden ser útiles para pensar estas preguntas y otras similares (no solo en el mencionado ámbito ético). En todo caso, vale la advertencia de que encontrarán aquí una reflexión de índo-

---

1 Cuando el Dr. Guillermo Lariguet me sumó a este proyecto de realizar un libro sobre Metaética, me entusiasmó volcar algunas de mis posiciones epistemológicas al campo disciplinar de la ética. Este tipo de trabajo conforma, de hecho, parte de nuestras conversaciones con el Prof. Lariguet, cuando se tratan de Filosofía.

2 La importancia de la ética con relación a su filosofía es confesada por el propio Wittgenstein (cf. Monk, 1997:178). Las razones por las que, a pesar de su importancia fundamental, su pensamiento en ética no se pone en palabras (T 6.42, 6.422), parten de la distinción lógica entre aquello que tiene sentido decir y aquello sobre lo que hay que callar, según la muy conocida observación del *Tractatus*. La actitud ética no puede, por ende y según Wittgenstein, proveer teorías morales, centrarse en la noción de acción o apoyarse en la idea de libertad. Esta posición deriva, además, de la noción de «independencia lógica» que los hechos tienen que sostener en última instancia (recordando que las figuras que nos hacemos de la realidad también son hechos), por lo que no hay nexo causal último entre los deseos, intenciones, etc., y lo que acaece en el mundo. Para una lectura en esta línea, consultar Tejedor (2015:138–155). En *Conferencia sobre ética* (CE:519), presentada en 1929, Wittgenstein se permite hablar un poco más, sugiriendo que esta mirada contingente de la realidad permite maravillarnos sobre el suceso de los hechos, que podrían no darse, pero de hecho se dan. *Investigaciones Filosóficas*, por su parte, solo menciona en dos pasajes a la ética (IF §77 y §304). Para una breve descripción del desarrollo del pensamiento ético de Wittgenstein, se puede consultar Diamond (1992).

le epistemológica y que no me adentraré en autores canónicos de la ética, por lo que me disculpo de antemano por cualquier expectativa que algún lector desprevenido pudiera aún guardar respecto al contenido de este trabajo.

### **Una perspectiva comunitarista del campo de la ética**

En un artículo titulado *Sobre la distinción entre Ética y Moral*, Gustavo Ortiz Millán (2016) toma como punto de partida la idea de que la Ética «se refiere a la afirmación de la conciencia individual autónoma o auténtica» (2016:113), en oposición a la «Moral», vinculada a la esfera de la observancia de reglas impuestas por la sociedad. Así, mientras sobre las costumbres no podemos opinar demasiado, la conducta ética es superior a la moral y sobre ella debemos ser más exigentes y prescriptivos: «el comportamiento ético, concebido como la afirmación de la conciencia autónoma y la autenticidad, debe tener un mayor peso que el ámbito de lo moral, entendido como la esfera de la observancia de reglas que nos son impuestas por la sociedad» (Ortiz Millán, 2016:115). Sobre esta base, el autor rastrea etimológicamente las expresiones «ética» y «moral» y pone en duda que la delimitación sea necesaria, llamando la atención sobre su carácter estipulativo y, por ende, revisable.

Mi intención en este trabajo es profundizar lo que interpreto de esta dirección de Ortiz Millán. Las caracterizaciones mencionadas de la «ética» y la «moral» apuntan según el escenario dicotómico inicial a la delimitación de dominios de objeto diferentes que requieren de dos capacidades distintas para su conocimiento. Este tipo de caracterización ha sido designada por Daniel Quesada (1998:18–19) en términos de una concepción metafísica del conocimiento, donde el dominio de objetos es el que determina los conceptos que puedo aplicar a través de capacidades que son específicas para dicho dominio. De manera análoga, podríamos decir que una «concepción metafísica de la ética» implica partir de la idea de que hay un dominio determinado de objetos éticos (que son diferentes a los objetos morales) y que sirven de razón para emplear ciertas capacidades individuales especiales para su captación (para captar lo que es bueno, o correcto, por ejemplo).

Una posición crítica a esta concepción metafísica o platónica,<sup>3</sup> entonces, puede impugnar este orden metafísico para pensar el campo o dominio de la ética. No solo se trata de analizar la corrección con la que empleamos determinadas palabras o si es posible reunir a la ética y a la moral bajo un término «vago» que permita emplearlo de una y otra manera según el caso. Implica

---

3 Quesada ilustra esta concepción metafísica con un pasaje de *República* (477b–478b).

pensar si estos dominios, caracterizados por las regiones separadas de lo individual y lo social, se hallan tan alejados como para demandar dos capacidades diferentes para su estudio y dos clases de conocimientos de naturaleza distinta como su resultado.

Los aportes de Wittgenstein en *Investigaciones Filosóficas*, respecto a la crítica de una concepción esencialista del significado (IF §§1-33), sirven de apoyo para el desacuerdo respecto de la concepción metafísica del campo de la ética. El lenguaje no tiene una sola función (IF §1, §2, §12, §14) originada en la aplicación de una definición ostensiva (IF §§28-35, §43, §68, §171) que determine de antemano el significado de las palabras a través del señalamiento de su objeto (IF §15, §40, §42). Comprender el significado de una palabra consiste en actuar o reaccionar del modo apropiado en relación con ella (más que de pensar de cierto modo) y esta es una conclusión a la que se arriba luego de haber objetado que cualquier señalamiento ostensivo puede ser malinterpretado (o interpretado de maneras diversas) no solo para el caso de los términos generales sino también de nombres propios. En otras palabras, la ostensión, lejos de su propósito de señalar de manera precisa el objeto, imparte una vaguedad al significado de las palabras. Ninguna definición ostensiva va a enseñarte cómo se usa la figura del rey en el ajedrez (IF §31): la definición ostensiva solo puede jugar un rol muy pequeño en el aprendizaje del significado de las palabras. Esta conclusión lleva a considerar que el supuesto metafísico que afirma que el significado está basado en el objeto es incorrecto. Según Wittgenstein, el significado de una palabra no es un objeto. El significado de una palabra es el modo en que la palabra es empleada en el lenguaje.

Ahora bien, si recogiendo esta posición nos trasladamos al campo de la ética, podríamos decir que esta no necesita una región propia para el empleo de sus conceptos, menos aún si se trata de una región de límites nítidos establecidos de antemano, sino que los emplea y los nutre de significado aplicándolos a la amplia esfera de las acciones humanas<sup>4</sup> o a la noción más nutrida de «forma de vida», entendida como el entramado de acuerdos, acciones y aspectos vitales en que estos se entrelazan y adquieren significado en nuestra vida. Esta forma de vida es el campo propio de la ética, el punto de partida donde el mundo y los seres humanos se encuentran entretejidos de una manera concreta y continúa, interesada. En esta forma de vida, además, lo individual y lo social se encuentran conectados, mis intenciones determinan mis empleos de la misma manera que el adiestramiento que he recibido al intro-

---

4 Y no solo a la esfera de las acciones entendidas en términos de conductas humanas, si pensamos en la perspectiva ofrecida por Wittgenstein en la *Conferencia sobre ética* (CE:519), la cual invita a una transformación en la manera en que experimentamos el mundo, como dotados de hechos que se dan pero que debido a su contingencia bien podrían no hacerlo.

ducirme en un juego de lenguaje normativizado, la familiaridad y la reacción proporcionada al contexto particular.

La adopción de una perspectiva wittgensteiniana en la crítica a la concepción platonista de las reglas como determinación de los significados se presenta al menos como un desafío al que una concepción universalista y apriorista de las normas éticas debe atender. La ventaja de la posición de Wittgenstein es que la ética y los valores, principios, hechos, juicios, etc., dejan de ser algo misterioso, sea por su origen o por su naturaleza. El desafío que esta conclusión impone, sin embargo, es que pareciera que no podemos decidir de antemano sobre qué versará la ética o qué analizará en cualquier circunstancia que se le presente. El término «bueno» no puede determinar por sí solo su aplicación,<sup>5</sup> dado que es posible pensar una interpretación o empleo que se corresponda con cualquier estado de cosas figurado. Además, la aplicación del concepto según algún criterio demandaría un nuevo criterio, que nos permita dirimir si el escogido en primer término ha sido el adecuado. Esto conlleva un regreso infinito en las interpretaciones y, por ende, la ausencia de un criterio último de corrección.<sup>6</sup> Entonces, podría ser que mi empleo, al no tener con qué medirse, sea en cada caso lo que considere (sea bueno, por qué no).<sup>7</sup>

---

5 En *The Varieties of Goodness*, Georg H. von Wright (1963:6–11) realiza una taxonomía de la multiplicidad de usos que puede tener la palabra «bueno». El primer caso es el de lo «bueno instrumental», por ej., cuando hablamos de un buen perro. También decimos que una persona es buena en algo (un buen científico). Von Wright le reserva el nombre de «bondad técnica» para resaltar un talento natural de una persona para llevar adelante determinada actividad. Otro caso es la «bondad médica»: un buen corazón, una buena memoria. Lo «bueno beneficioso» resume ejemplos como el que tener buenas instituciones es bueno para todos quienes habitan dicho país. Lo «bueno hedonista» es aquello que trae placer: un buen chiste, un buen vino, una buena compañía. También están los empleos emparentados de modo más directo con la moral, y son aquellos que tienen que ver con la conducta y el carácter de los seres humanos, con sus motivos e intenciones. Estos ejemplos muestran la «multiplicidad semántica y la riqueza lógica» del fenómeno, siendo que además, según aclara von Wright, estas formas de empleos no pueden ser reunidas bajo un único «bien genérico» de las cuales podrían considerarse especies ni tampoco deberse a una acepción «vaga» del término.

6 Este es un problema que desde luego también mantienen las concepciones platonistas de las que hemos tomado distancia de la mano de Wittgenstein. Una definición universal para lo que es bueno, una suerte de regla que contenga todos los criterios posibles para su aplicación de forma a priori lleva a preguntar ¿cómo haríamos para escoger el criterio adecuado entre todos estos criterios, de forma a priori?

7 El problema del no-cognitvismo es similar al planteado por la paradoja wittgensteiniana: nadie puede decirte qué hacer, qué es lo correcto y qué no, dado que no hay ningún criterio objetivo sobre el que sostenerse. Piénsese si no, en la afirmación de Hare (1963:2) según la cual cada individuo tiene la libertad de formar sus propias opiniones acerca de cuestiones morales y que esta libertad de tomar posición es mucho más libre, menos constreñida (para evitar la redundancia) en tanto no es limitada por la realidad, los hechos. Esta es una consecuencia que planteada en los extremos que formulamos resulta evidentemente indeseable para el campo de la

La atención a este desafío debe incluir, por lo tanto, algún rasgo que permita reintroducir el carácter normativo de la regla, que nos permita contar con un estándar para «fijar» el significado de un concepto. Una manera de atender a esta demanda sin comprometerse con una perspectiva esencialista para la determinación del significado de conceptos (incluyendo los vinculados a atribuciones morales como «bueno», etc.) es afirmar que un comportamiento particular es o no ético, por ejemplo, dependiendo de si conforma o entra en conflicto con los estándares diversos que se sostienen de manera comunitaria.

La solución propuesta por Kripke (1989 – parte II) a la paradoja del seguimiento de reglas es muy conocida y apunta en esta dirección comunitaria que deseo poner en juego. Las condiciones de asertabilidad o la justificación en el empleo del lenguaje con un significado determinado están dadas por la comunidad (Kripke, 1989:78). Si nos limitamos a ver aisladamente a una sola persona (sus estados psicológicos y su conducta), todo lo que podemos decir es que está autorizada a aplicar la regla del modo en que se le ocurra. Esto no soluciona el problema escéptico: «pensar que uno está obedeciendo una regla no es obedecer una regla» (IF §202). De manera similar en el campo de la ética, pensar que uno está actuando correctamente no garantiza que uno esté actuando correctamente. Esto se soluciona si consideramos que la comunidad a la que la persona pertenece aporta las condiciones de justificación para la predicación de «bueno» y similares. Según las palabras de Kripke:

La situación es muy diferente si ampliamos nuestra perspectiva al considerar a quien sigue la regla y nos permitimos considerarlo interactuando con una comunidad más amplia. Entonces otros tendrán condiciones de justificación para predicar del sujeto si sigue la regla correcta o incorrectamente y éstas no serán sencillamente que la propia autoridad del sujeto ha de aceptarse incondicionalmente. (1989:87)

La manera en que la comunidad aporta este criterio de corrección es explicada por Kripke a partir del «juego de atribución de conceptos» (1989:92), según el cual se dice que una persona domina un concepto (lo comprende) si la mayoría de sus respuestas o empleos particulares del mismo concuerdan (en cantidad suficiente) con los de la comunidad a la que pertenece. Por ejem-

---

ética. Wittgenstein, por su parte, identifica este tipo de problema semántico en la paradoja del seguimiento de reglas, la cual expresa de la siguiente manera: «una regla no podía determinar ningún curso de acción porque todo curso de acción puede hacerse concordar con la regla. La respuesta era: si todo puede hacerse concordar con la regla, entonces también puede hacerse discordar. De donde no habría ni concordancia ni desacuerdo» (IF §201).



plo, para considerar que una persona emprendió una acción buena desde el punto de vista ético, es necesario que domine este tipo de comportamientos, que cumpla con las expectativas que de él tenemos y que se lo sancione e incluso excluya si no cumple con lo esperado en numerosas ocasiones. Es decir, la consideración de actos particulares de un individuo en términos éticos se apoya en las prácticas en las que una comunidad concuerda, en sus criterios, en su forma de vida.

No quiero afirmar con esto que el campo de la ética no está sembrado de reglas. Si el significado que tiene un término moral como «bueno» es el uso en el lenguaje y no la regla que gobierna el uso, esto es algo que puede sostenerse incluso si pensamos en la ética como un área gobernada por reglas exhaustivas.<sup>8</sup> Quisiera intentar ser más clara respecto de este tema. Wittgenstein distingue entre áreas del lenguaje que no están gobernadas por reglas (*IF* §81, §83) y otras, como es el caso de la «adición» y otros ejemplos de matemáticas que incluso estando gobernadas por reglas rígidas, no deben ser malinterpretadas en términos mentalistas. El problema reside, entonces, en pensar a las reglas como un signo mental privado, algo que aparece delante de nuestra mente para que comprendamos el significado de «bueno», por ejemplo. Es decir, podemos contar con la noción de regla para determinar lo que es bueno, pero el uso es central para comprender el significado ya que la captación y la internalización de signos no pueden cumplir la función de decirnos cuándo empleamos correctamente la palabra «bueno» para describir una acción de este tipo. Las razones por las que no puede hacerlo son las esgrimidas al comienzo de este apartado: el problema de la concepción mentalista de las reglas es que implica que comprender cualquier regla dada requiere comprender una serie infinita de otras reglas. Perdemos el criterio de corrección.

No estoy en condiciones de decir cuál es la regla o reglas que gobiernan el empleo de las propiedades morales que se incluyen en el campo de la ética. Asumo que aquí habrá múltiples posiciones rivales y cualquier intento mío sería vano. Pero podemos suponer, si queremos asumir que «bueno», en tanto pertenece a una disciplina de estudio como la ética, no es un concepto vago, debe comportarse de manera similar a como se comportan conceptos como

---

8 Algunas lecturas wittgensteinianas afirman que no puede haber una metaética porque no hay un criterio de demarcación de la ética, porque no se posee un criterio rígido que determine de antemano qué cae bajo su dominio y qué no. Esta posición ha sido sostenida, por ejemplo, por Stephen Mulhall (2002:703). Esto trae la indeseable consecuencia de que no hay un vocabulario moral ni una disciplina ética, con lo cual no tenemos modo alguno de decir que algo es malo o bueno con sentido. Pero creo que el problema es aún más radical, desde que lecturas como las de Mulhall sostienen que «ética» es un concepto vago, pero el problema persistiría aun si consideramos que «ética» es, después de todo, un concepto rígido.

«adición» en aritmética. En otras palabras, podríamos asumir simplemente que es posible lograr una suerte de regla (o conjunto de reglas) prescriptivas a priori que nos digan qué hacer en casos en los que nos hallemos ante la aplicación de, por ejemplo, el término «bueno».<sup>9</sup>

Entonces podríamos decir, por ejemplo, que ético resulta todo lo que es bueno (y aquí damos una regla rígida), pero lo que es bueno dependerá de las circunstancias concretas en que se desarrolle la acción, del contexto y las intenciones del agente que las desarrolle, de su educación o adiestramiento en el seno de una comunidad en la que participa del juego de atribución de conceptos, así como de sus habilidades y destrezas adquiridas, etc. Esto tal vez nos ayude a ser más comprensivos respecto de las acciones éticas de las personas. Cuando digo «ser comprensivos», me refiero precisamente a la consideración de estos aspectos externos a las propias intenciones o hechos mentales ostentados por el agente moral, dotando a la expresión de un tinte, nuevamente, wittgensteiniano. Entonces, cuando digo que nos ayuda a ser más comprensivos no estoy diciendo que nos ayudará a ser más compasivos (ese vínculo merece un estudio aparte y sospecho que no es tan directa su unión). Estoy diciendo que la consideración de los factores externos contribuye a la decisión de si una persona que realiza una acción particular en una comunidad en la que pertenece está comportándose de manera ética o no.

En conclusión, la regla que determina, por caso, lo que es «bueno», puede ser rígida, pero su empleo siempre está indeterminado si se lo considera de forma a priori. La única manera de tener un criterio sobre el que decir que determinada acción es correcta o incorrecta, es contar con la sanción comunitaria respecto a los comportamientos significativos de cada individuo. Este modo de pensar, creo yo, hace que la ética deje de delinarse como una disciplina eminentemente individual, sin por ello abandonar la idea de un agente distinguible dentro de la comunidad.

Por ende, no parece necesario comprometerse con una concepción esencialista que determine de antemano todos los casos posibles en que podemos emplear el término «bueno» en referencia a un hecho. Pero tampoco parece necesario asumir el otro extremo del arco, una conclusión escéptica que nos diga que nunca comprendemos realmente el significado de algo excepto que busquemos una solución estipulativa consensualista y nos conformemos con ello. Lo que hacemos, en todo caso, es comportarnos de manera directa (*IF* §201, §202), actuar de acuerdo con una práctica de la que somos parte (*RFM*, VI §43). En palabras de Wittgenstein: «esta no es una concordancia de opiniones, sino de forma de vida» (*IF* §241). De otra manera, colapsa la distin-

---

9 El imperativo categórico kantiano podría aquí servir de ejemplo.

ción entre «correcto» e «incorrecto». Para responder al desafío escéptico que dice que no hay ningún hecho al que podamos recurrir para decir qué es lo que denota un concepto, no solo tengo dos alternativas: considerar la versión platonista de los hechos o asumir la solución escéptica<sup>10</sup> de Kripke que afirma que los hechos que demanda el escéptico no son necesarios.

## El punto justo de los hechos de la ética

La lectura kripkeana del seguimiento de reglas, al igual que la posición de Glaucón<sup>11</sup> en *República* (quien consideraba a la moral como una regla que las personas introducen para garantizar la paz y la estabilidad social), han sido catalogadas como versiones convencionalistas o consensualistas en sus respectivos campos, enfrentando los ataques del caso. No reseñaré estas críticas<sup>12</sup> ni intentaré tampoco medir mi posición respecto de ellas. Lo que quisiera señalar, en cambio, es que la lectura comunitarista propuesta en el apartado anterior no conlleva necesariamente la presunción de que la regla que determina si algo es bueno, por ejemplo, dejaría de existir y que, por ende, no contamos con ningún criterio objetivo para medir nuestras acciones. Lo que mi posición afirma, a diferencia de lo recién dicho, es que las reglas que determinan de manera a priori y universalmente no pueden sostenerse por sí mismas, porque las reglas existen en tanto tengan o adquieran un significado en nuestras vidas, mientras se sostengan en nuestra forma de vida, porque cumplen un rol en nuestra «vida civil» (*IF* §125) y allí adquieren su importancia. Y esto es así para el caso de conceptos morales como respecto de reglas rígidas como las que gobiernan la «adición».

La diferencia que planteo respecto de la solución escéptica a la paradoja wittgensteiniana sostenida por Kripke, consiste en asumir que hay hechos morales, que es factible considerar su existencia, aunque estos hechos dejen de ser crudos, sin rastro humano alguno.<sup>13</sup> Los hechos están nutridos por el entramado de instituciones, contextos y aspectos familiares que conforman nuestra forma de vida. No son sustancias separadas entre las que hubiera que

---

10 Kripke denomina a su salida a la paradoja una «solución escéptica», porque reconoce que las afirmaciones negativas del escéptico son incontestables, pero señala que nuestras prácticas en cuestión no necesitan la justificación que el escéptico demanda.

11 Mencionada al comienzo de este trabajo.

12 Para una versión de las críticas a las posiciones consensualistas en ética, ver Sayre—McCard (2014). Podemos encontrar ejemplos de crítica a la lectura comunitarista de Kripke en Canfield (1996) y Hacker (2010).

13 Para una discusión ampliada de diversos sentidos del sintagma «hecho moral» puede verse el capítulo titulado «Realismo moral» de Guillermo Lariguet, en esta misma obra.

escoger a fin de ser un realista metafísico o un convencionalista. Esta es, desde mi perspectiva wittgensteiniana, una falsa dicotomía, un seudoproblema. Los hechos no son plenamente subjetivos ni completamente objetivos, independientes.

Lo que quiero manifestar es que la conclusión que Kripke saca de su lectura de Wittgenstein no es compulsiva en lo que respecta al componente escéptico que consiste en afirmar que no hay hechos. *Realism and the Realistic Spirit* de Cora Diamond (1991)<sup>14</sup> es un claro ejemplo (anticipado, dado que su posición es anterior al planteo de Kripke) de cómo se puede realizar una lectura similar respecto al diagnóstico de la paradoja escéptica escogiendo otra vía para su solución. No es necesario eliminar los hechos para introducir la vida civil, al individuo viviendo en la comunidad.

Resulta notorio de esta comparación entre las alternativas interpretativas de Kripke y Diamond el que ambos enfatizan el aire berkeleyano de la posición de Wittgenstein, aunque en direcciones opuestas respecto a sus implicancias. Según la reconstrucción de Kripke (1999:68–69), Berkeley sostiene que la afirmación de la materia y de lo extramental son el producto de una interpretación metafísica errónea de nuestras creencias comunes. Wittgenstein, por su parte, sostiene que el intento de encontrar un hecho superlativo (*IF* §192) en mi mente que constituya el referirme a la adición mediante «más», y que determine de antemano qué debería hacer para concordar con este hecho, está basado en una construcción filosófica errada de expresiones ordinarias como «quiso decir tal y tal cosa», y otras (*IF* §§183–93).

Cora Diamond (1974:82), en cambio, sostiene que Wittgenstein no es un escéptico respecto de la materia y lo extramental, sino que es un realista no metafísico, un realista en el sentido «civil» del término. La afirmación de Wittgenstein que aparece en *Remarks on the Foundations of Mathematics* (*RFM* VI, §23), a saber: «No empirismo y aun así realismo. Esto es lo difícil», nuclea gran parte de las discusiones acerca del realismo en la propuesta filosófica del segundo Wittgenstein y es sin duda el punto central en torno al cual gira el mencionado artículo de Diamond (1991). El realismo al que se refiere Diamond es «un punto de vista en el cual de una u otra forma se enfatiza la significancia de lo que es independiente de nuestro pensamiento y experiencia» (1991:39). Esta es sin dudas una caracterización general y básica dentro de la cual deberemos colocar a una posición a fin de llamarla «realista» pero que amerita, además, otros añadidos en cuanto a su determinación a fin de poder identificar de manera correcta a qué clase o variante de realismo nos estamos refiriendo. En sus consideraciones, la autora sostiene que puede notarse un

---

14 Hilary Putnam (1998) puede ser otro ejemplo de este tipo de posición realista que sostengo.

«espíritu», un parecido de familia podríamos decir, entre la posición de Wittgenstein y el realismo al enfatizar que el significado es independiente de cada uno de nosotros, pero no se trata de una concepción «metafísica» o esencialista ya que lo que asumimos, en realidad, es un llamado a ser «realistas» en la acepción vulgar del término, es decir, atendiendo a los casos y contextos concretos en vez de abstrayéndonos de ellos. La realidad queda aquí configurada como algo interno a la propia vida cotidiana, una vida de un sujeto que no está «sin mundo», sin constricciones, sino que es una forma de vida que se desenvuelve dentro de un plano ya dado, contextualizado, de significados compartidos y objetivados a los cuales debemos responder.

En esta nueva caracterización, hay toda una familia de empleos no filosóficos de la expresión «realismo» que salen a la luz. Diamond (1991:40-41) señala tres ejemplos extraños a la filosofía: 1) Decimos que alguien es realista cuando está sosteniendo algo que está de acuerdo con los hechos y que no lo es, consecuentemente, en caso contrario o cuando se niega a mirar los hechos. 2) Las novelas e historias también pueden ser o no realistas dependiendo la atención que ellas presten al detalle y a la particularidad. Y las ficciones pueden ser realistas, aunque evidentemente no reproduzcan todo o sean una simplificación, aunque es evidente que ciertas cosas no ocurrirán en ellas: las macetas no hablan, las personas no regresan en el tiempo, los elfos no les hablan a zapateros, etc. 3) Por último, este aire realista le otorga relevancia a la causalidad: en una historia realista los eventos se desenvuelven unos a otros y los personajes responden a las circunstancias. En ella se encuentra en marcha el modo como funcionan las cosas en nuestra vida: qué conduce a qué, qué tipo de cosas determinan los hechos concretos como procedieron los eventos.

Estos ejemplos nos muestran que probablemente en la acepción que Wittgenstein ofrecía para defender una postura realista cabe un lugar para el empirismo, aunque en sentido modificado. Se trata de la experiencia de nuestra «vida civil», de nuestros comportamientos cotidianos y compartidos, ateniéndose a normas de empleo ya instauradas, la que conforma la realidad con la que debemos mantenernos en contacto y coherencia si queremos ser realistas. No se trata de una realidad desnuda y completamente ajena a nuestras experiencias, ni una experiencia de origen meramente perceptiva, subjetiva e individual o convencional, la que sirve como criterio de corrección de nuestros comportamientos lingüísticos. Este es un realismo que evita el empirismo reduccionista y puede quedarse apegado, sin embargo, a su aplicación a los hechos. En un pasaje de la compilación que va de los años 1942-1944 de *Remarks on Foundations of Mathematics (RFM)*, Wittgenstein dice:

¿Calcula la calculadora? Imagina que hubiera aparecido por casualidad una calculadora; y que alguien aprieta casualmente sus botones (o un animal anda sobre ellos) y realiza el producto  $25 \times 20$ . Quiero decir: es esencial a la matemática que sus signos se usen también en lo civil [*mufti*]. Es el uso fuera de la matemática, es decir, el significado de los signos, lo que convierte en matemática el juego de signos. (*RFM* v, §2)

Wittgenstein es un realista en tanto presta atención a las matemáticas en sus empleos y actividades ajenos a la matemática. El uso de la matemática en tales actividades es la realidad ordinaria que la sostiene en su lugar y que la salva de ser transformada en una forma arbitraria de los símbolos. Haciendo extensión de esta analogía que he venido manteniendo entre conceptos y reglas matemáticas y éticas, creo que es posible pensar en una suerte de hechos éticos civiles, que tengan un sentido similar a los desarrollados por Diamond. No quiero decir que entonces la ética (y la moral, que se vio arrastrada aquí de manera forzosa) pueda tratar sobre peras y sillas, porque en realidad puede tratar sobre cualquier cosa, por no tener un campo al que se aplique por su propia definición. Lo que estoy diciendo es que puede tratar sobre «cualquier cosa» en la medida en que esas mismas cosas tengan una importancia o rol significativo en nuestras vidas, sea que decidamos aplicarle propiedades morales como «bueno» o que expresemos un significado moral a través de un gesto, o de la narración de una historia.

La «realidad» no es otra cosa que «el terreno áspero» (*IF* §107), el modo en que utilizamos las palabras, en qué contextos lo aplicamos, las consideraciones que realizamos en nuestra elección de ellas, la educación que hemos recibido, etc. Wittgenstein afirma en *Investigaciones Filosóficas*: «Pregúntate siempre en esta dificultad: ¿cómo hemos aprendido el significado de esta palabra (“bueno”, por ejemplo)? ¿A partir de qué ejemplos; en qué juegos de lenguaje? Verás entonces fácilmente que la palabra ha de tener una familia de significados» (*IF* §77). Las reglas de «adición», o las que dicen cómo clasificar algo según su bondad, solo adquieren significado y relevancia en nuestras vidas, y son los empleos que de ellos hacemos, en el seno de una comunidad que les atribuye significado a los conceptos y también en la atención a los contextos, a mis propias intenciones y a lo familiar que me resulten las cosas, las que determinan cuál es su significado.

Según lo que he estado expresando, Diamond, pero también Putnam y McDowell, respaldan la afirmación de que no es necesario desprenderse de los hechos morales y de que es posible sostener (con mayor o menor intensidad) un compromiso realista en la determinación del significado de conceptos (o en el seguimiento de reglas). Lo que quiero respaldar ahora, brevemente, es la

idea de que este tipo de posición realista no reductivista y comunitarista incluida en la consideración del seguimiento de reglas tiene también un lugar cuando pensamos en los hechos morales.

Como ejemplo aproximado de esta posibilidad de extender esta posición wittgensteiniana al campo de la ética, podemos citar la tesis sostenida en *Realism and Imagination in Ethics* de Sabina Lovibond (1983). Esta autora defiende una forma de realismo moral derivada de la concepción del lenguaje del segundo Wittgenstein, recogiendo su mencionada afirmación sobre el desafío de ser realista sin ser empirista (en el sentido representacional y reductivista del término). Su apuesta por un realismo moral dentro del marco wittgensteiniano le permite tomar distancia del no-cognitismo, entendido como la tesis de que no hay cognición o conocimiento moral por no haber hechos morales objetivos ni posibilidad de establecer enunciados morales verdaderos en términos de correspondencia con dichos hechos. Lovibond aporta argumentos a favor de «hechos morales» consistentes en «una circunstancia registrada por un enunciado indicativo que puede ser usado de manera simultánea para realizar reportes sobre el mundo y para proveer de una justificación racional no hipotética de las acciones» (1983:23). Además, propone una clase de percepciones puramente cognitivas de estos hechos.

En conclusión, los hechos de la ética no están crudos, pero tampoco completamente cocidos. La ética puede ocuparse del «estado civil» (*bürgerliche Stellung*, *IF* §125) de los hechos que consideramos «buenos», y de por qué lo son. Ese mundo civil podría ser el objeto del problema ético, más allá de donde no se podría excavar, porque no tendría sentido hacerlo.<sup>15</sup>

## Consideraciones finales

Al comienzo de este trabajo presenté dos ideas a defender, las cuales pueden ser ahora planteadas bajo la forma de preguntas a la luz de las cuales medir los aportes realizados. En primer lugar, podemos cuestionarnos si la ética versa efectivamente sobre una dimensión individual de los seres humanos. En segundo lugar, tenemos que preguntar si hay lugar para los hechos de la ética. Ambas preguntas responden al intento de comprender los presupuestos y compromisos del campo disciplinar de la ética, de las categorías con las que trabaja y de su práctica.

---

<sup>15</sup> En esta receta de cocción, restaría aclarar el rol de las emociones, la educación y el impacto de los conceptos en lo que percibimos como hechos morales.

Las respuestas a las que he arribado pueden resumirse de la siguiente manera: en relación con la primera pregunta, las regiones de lo individual y lo social no se hallan tan separadas como para demandar dos capacidades diferentes para su estudio y dos clases de conocimientos de naturaleza distinta como su resultado. En este sentido, no es necesario sostener una perspectiva metafísica que determine el campo de estudio de la ética. Lo individual y lo social se encuentran conectados, además, dado que la postulación de reglas universales y a priori para la determinación del significado de propiedades morales como «bueno» se ha mostrado como un camino al menos dificultoso si se desea sostener un criterio de corrección para los mismos. Lo que propuse, en cambio, es considerar que mis intenciones determinan mis empleos de la misma manera que el adiestramiento que he recibido al introducirme en un juego de lenguaje normativizado, la familiaridad y la reacción proporcionada al contexto particular. En este sentido, mis acciones (incluyendo preferencias) éticas están constreñidas por una comunidad en la que participo y que sostiene un juego de atribución de conceptos, regulando los empleos. De esta manera, la ética deja de delinearse como una disciplina eminentemente individual, sin por ello abandonar la idea de un agente distinguible dentro de la comunidad.

El segundo interrogante que tracé para este trabajo puede desprenderse de las consideraciones vinculadas a la primera respuesta que esgrimí, dado que si no resulta necesario comprometerse con una concepción esencialista que determine de antemano todos los casos posibles en que podemos emplear un mismo término moral en referencia a un hecho y si la comunidad debe participar del establecimiento de los criterios de corrección para el empleo de los mismos, entonces tal vez no haya hechos a los que recurrir en ética. Según mi punto de vista, las reglas que determinan de manera a priori y universalmente no pueden sostenerse por sí mismas, sino que existen en tanto tengan o adquieran un significado en nuestras vidas. Las reglas se sostienen en nuestra forma de vida porque son relevantes en nuestra «vida civil». Sin embargo, aunque no haya objetos platónicos, podemos asumir que hay hechos morales que habitan un mundo humano común, compartido. Los hechos están nutridos por el entramado de instituciones, contextos y aspectos familiares que conforman nuestra forma de vida. Esta posición, además, puede enmarcarse en cierta acepción corriente de la palabra «realismo», tal como es delineada a través de algunos ejemplos por Diamond (1991). Esta línea permite entender que no es necesario eliminar los hechos para introducir la vida civil de un individuo y que tampoco es necesario reducir los hechos morales a un conjunto de características objetivas del mundo, en el sentido de ser completamente independientes de los sujetos, a fin de asegurar un criterio de corrección para las acciones morales.



El enfoque wittgensteiniano que adopté en el tratamiento de estas cuestiones parte de la consideración de que las propias discusiones sobre la existencia y naturaleza de los hechos morales o aquellas vinculadas a los estándares morales y su vínculo con la cultura presentan una importante conexión con algunas de las preocupaciones filosóficas de Wittgenstein, recogidas y desarrolladas en direcciones diversas por algunos de sus intérpretes. Por ejemplo, la pregunta «¿cómo es que los hechos morales establecen un estándar apropiado para nuestros comportamientos?» es del todo análoga a «¿cómo es que en cada caso sabemos que seguimos una regla determinada de la manera correcta?».

Este tipo de posición epistemológica respecto de la Ética mantiene, además, una perspectiva social. Ella se compromete con la idea de que el conocimiento moral existe, pero asume que tales hechos morales son relativos a aquellos que comparten una forma de vida (para decirlo en términos más generales o menos wittgensteinianos, para un mismo grupo social). Esto, por supuesto, no solo se traduce en críticas a posiciones aprioristas o universalistas, sino que a la par proporciona un punto de arranque para pensar los desacuerdos y la diversidad morales sin caer en posiciones relativistas e individualistas o escépticas (que afirmen que ningún hecho moral existe y por ende no hay nada que pueda ser conocido).

Pareciera que estas posiciones que he estado delineando nos conducen a un naturalismo en ética, con el riesgo de caer en la falacia naturalista humeana (Hume, 1984 – L.III, P. I, Sec. I). No es de buen gusto comenzar una nueva discusión en el apartado de las conclusiones, por lo que solo haré un breve comentario. He identificado las propiedades morales con aspectos naturales de nuestro mundo, pero no he reducido estos aspectos al sustrato último de la ciencia naturales sino, antes bien, a nosotros que estamos en este mundo. La identificación de las propiedades morales con un entramado natural de cuestiones (la educación, el contexto particular, la comunidad, etc.) implica, si se quiere, un naturalismo laxo pero que de ningún modo reduce el carácter normativo de las reglas a un componente meramente descriptivo de las ciencias empíricas. En este sentido, debería tal vez dar respuesta al interrogante fundamental respecto de si mi posición logra captar o no la finalidad ética que consiste en decidir cuándo algo está bien y cuándo está mal. Bueno, lo que tengo para responder por el momento es que tal vez haya que abandonar la idea de que Dios diferencia el bien y el mal y nosotros solo tenemos que seguirlo. Esta sería una falacia que cobra la forma de un anhelo normativista: el intento de derivar lo que es, de lo que debería ser el bien como norma prescriptiva absoluta.

Creo que Wittgenstein mantuvo a lo largo de sus períodos de pensamiento este espíritu que consiste en afirmar que la ética no es una teoría sino y, ante

todo, una forma de vida, algo que se muestra en nuestros comportamientos y en el modo en que tenemos de experimentar lo que nos rodea. No podemos decir lo que es «bueno» en abstracto,<sup>16</sup> sin considerar los hechos, aunque el *Tractatus* y las *Investigaciones Filosóficas* realizan un recorrido diferente para arribar a esta conclusión. En el caso del *Tractatus*, solo la asunción de la radical independencia lógica de los hechos atómicos (incluyendo nuestros pensamientos y proposiciones, que también son hechos), puede conducirnos a la transformación ética en el modo en que nos paramos en relación con ellos, asumiendo su contingencia y maravillándonos de su ocurrencia. En el caso de las *Investigaciones Filosóficas*, aunque contemos con una regla que determine lo que es «bueno», el concepto solo determina o completa su significado cuando se lo emplea en un juego de lenguaje comunitario, así como en la consideración del contexto y de la intención del hablante, además de la familiaridad y la práctica compartida que supone su uso.

---

16 En sus *Principia Ethica* (1993), Moore consideraba que de lo bueno, en abstracto, no se puede decir nada. Solo puede hablarse de la bondad cuando se adjudica como una propiedad a algo en concreto. Algo de esto también está en la inspiración de Wittgenstein al colocar a la ética dentro del campo de lo que no se puede hablar en el *Tractatus*. Esta línea de lectura podría, además, colocar al *Tractatus* en la misma línea crítica del naturalismo presentada por la «Open Question» de Moore (1993:95), cuya finalidad es mostrar que las explicaciones naturalistas de la moral están equivocadas.

## Referencias bibliográficas

- Canfield, J.V. (1996). The community view. *The Philosophical Review*, 105, 469–488.
- Diamond, C. (1991). Realism and the Realistic Spirit [1974–1982]. En Diamond, C., *The Realistic Spirit / Wittgenstein, Philosophy and the Mind* (pp. 39–71). MIT Press.
- Diamond, C. (1992). Ludwig Wittgenstein. En Becker, L.C. (Ed.), *Encyclopedia of Ethics* (pp. 1319–1324). Garland Pub.
- Diamond, C. (1996). Wittgenstein, Mathematics and Ethics. Resisting the Attractions of Realism. En Stern, D. y Sluga, H. (Eds.), *The Cambridge Companion to Wittgenstein* (pp. 226–260). Cambridge.
- Dworkin, R. (2011). *Justice for Hedgehogs*. Cambridge, Harvard U.P.
- Hacker, P.M.S. (2010). Robinson Crusoe Sails Again: The Interpretative Relevance of Wittgenstein's Nachlass. En Venturhina, N. (Ed.), *Wittgenstein After His Nachlass* (pp. 91–109). Palgrave Macmillan.
- Hare, R.M. (1963). *Freedom and Reason*. Oxford U.P.
- Hume, D. ([1739]1984). *Tratado de la naturaleza humana*. Tomo 3. Hyspamérica.
- Kripke, S. (1989). *Wittgenstein: Reglas y lenguaje privado*. UNAM (México).
- Lovibond, S. (1983). *Realism and Imagination in Ethics*. University of Minnesota Press.
- Monk, R. (1994). *Ludwig Wittgenstein. El deber de un genio*. Anagrama.
- Moore, G. E. ([1903]1993). *Principia Ethica*. Cambridge University Press.
- Mullhal, S. (2002). Ethics in the Light of Wittgenstein. *Philosophical Papers* (Special Issue), 31, 293–321.
- Ortiz Millán, G. (2016). Sobre la distinción entre ética y moral. *Isonomía*, (45), 113–139.
- Platón (2012). *República* (trad. de Antonio Camarero). Eudeba.
- Putnam, H. (1998). Kripkean Realism and Wittgenstein Realism. En Biletzki, A y Matar, A. (Eds.), *The Story of Analytic Philosophy* (pp. 241–252). Routledge.
- Quesada, D. (1998). *Saber, opinión y ciencia*. Ariel.
- Sayre–McCord, G. (2014). Metaethics. En Zalta, E., *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2014/entries/metaethics/>.
- Tejedor, C. (2015). *The Early Wittgenstein on metaphysics, natural science, language, and Value*. Routledge.
- Von Wright, G.H. (1963). *The Varieties of Goodness*. Routledge & Kegan Paul / The Humanities Press.
- Wittgenstein, L. (RFM) (1981). *Remarks on the Foundations of Mathematics* (1933/1944). En von Wright, G.H.; Rhees, R. y Anscombe, G.E.M. (Eds.). Blackwell (versión en español: *Observaciones sobre los fundamentos de la matemática* —trad. Isidoro Reguera—). Alianza.
- Wittgenstein, L. (CE) (2009). Conferencia sobre ética. En *Wittgenstein II* (pp. 513–524). Gredos.
- Wittgenstein, L. (IF) (2009). Investigaciones Filosóficas. En *Wittgenstein I*. Gredos.
- Wittgenstein, L. (T). *Tractatus Lógico-Philosophicus* (trad. de Jacobo Muñoz e Isidoro Reguera). Alianza.

## **20. «Eso no es gracioso»: acerca de cómo los chistes misóginos ayudan a comprender la naturaleza de los valores**

*Facundo García Valverde*

Nuestra vida como agentes prácticos es más amplia que nuestros juicios normativos sobre lo que debemos o no debemos hacer. En ella, nos enfrentamos con situaciones que nos resultan envidiables, temibles o cómicas; observamos acciones ajenas que son admirables, personas despreciables y con hechos fortuitos que son dignos de pena. En cada una de estas situaciones atribuimos un predicado evaluativo a un determinado objeto.

La pregunta acerca de la naturaleza de los valores es una pregunta eminentemente metaética: ¿qué son los valores?; ¿son la proyección de nuestros deseos e intereses?, ¿son elementos de la fábrica del mundo que producen determinadas respuestas en el sujeto práctico?, ¿son análogos a propiedades primarias (forma, etc.) o a propiedades secundarias (color, etc.)?; ¿qué relación existe entre las emociones y los valores? Como vemos, el problema de la naturaleza de los valores es multifacético y cada uno de sus rostros genera interrogantes propios no necesariamente vinculados entre sí. En una reformulación del *Eutifrón* platónico, podríamos preguntarnos con Risieri Frondizi (1958:19): ¿tienen valor las cosas porque las deseamos o las deseamos porque tienen valor?

Si bien este capítulo reconstruirá en términos generales el mapa de la discusión, se concentrará en un tipo especial de respuesta lo suficientemente promisorio como para ser analizada en detalle, la Concepción del Valor como Actitud Idónea (CVA). De acuerdo con ella, los valores deben ser analizados en términos de las respuestas evaluativas que son idóneas (*fitting*) para ese objeto; por ejemplo, que un policía con el arma reglamentaria desenfundada sea temible debe ser analizado en términos de si esa respuesta evaluativa (la de sentir temor) es adecuada para ese objeto y situación.

En la primera sección del capítulo, realizaremos una taxonomía simple de las concepciones acerca de la naturaleza de los valores, en la cual ubicaremos a CVA como superadora de los problemas de posiciones realistas e idealistas. En la segunda sección, reconstruiremos el principal problema metaético de CVA (el Problema del Entrevero), según el cual la noción de idoneidad no es suficiente para determinar cuándo un valor es atribuido por razones correctas y cuándo es atribuido por razones incorrectas. Ejemplificaremos este problema con los chistes misóginos cómicos, es decir, con casos donde la respuesta emocional idónea no se correspondería con las propiedades evaluativas del

objeto. En la tercera sección, evaluaremos algunas soluciones propuestas a este problema y trataremos de identificar el centro neurálgico de la incapacidad de las CVA para dar una respuesta adecuada. Como intentaremos mostrar, este centro neurálgico reside en sostener que las propiedades evaluativas tienen que estar relacionadas con una única consideración idónea y no con una lectura integral del contexto práctico.

## **La Concepción del Valor como Actitud Idónea**

A pesar de que CVA es una categoría reciente que se aplicó a un conjunto heterogéneo de posiciones sobre la naturaleza del valor (Rabinowicz y Rønnow-Rasmussen, 2004), puede ser caracterizada por la afirmación de dos tesis genéricas.

La primera de ellas es la tesis del carácter antropocéntrico de los valores, en el sentido de que el valor es indisoluble de las respuestas y actitudes humanas de evaluación; de acuerdo con ella, hay una conexión conceptual entre valores y actitudes, en especial, las emociones. Según esta conexión, estas respuestas y actitudes determinan parcialmente a los valores y, por lo tanto, los valores son dependientes de ellas (McDowell, 1998). Por ejemplo, David Wiggins afirmó que «x es bueno si y solo si x es la clase de cosas que hace apropiado o invoca cierto sentimiento de aprobación dado el rango de propensiones que tenemos de responder de esta u otra forma» (1987:206).

La segunda tesis compartida es la tesis de la prioridad deóntica, es decir, que el dominio axiológico se reduce al dominio normativo y, en específico, al dominio deóntico. Los valores son considerados como la respuesta o la actitud que *debe* tenerse frente al objeto en cuestión y, en este sentido, tales actitudes están constreñidas por consideraciones de idoneidad; que algo sea valioso no es explicado como el mero informe de las actitudes que los humanos (o una mayoría de ellos) tienen frente a ese objeto sino como la idoneidad de ciertas actitudes favorables (que los humanos tienen razón en tener) hacia ese objeto. Así, la actitud de admiración hacia *Morning Sun* de Edward Hopper es apropiada porque existe una serie de normas que dictan la idoneidad de la admiración hacia esa pintura. Puesto de forma negativa, cuando un agente afirma que *Morning Sun* es admirable está afirmando que no es apropiado no sentir admiración frente a esa pintura. Si bien Franz Brentano y Henry Sidgwick son considerados los antecesores de esta segunda tesis (Rabinowicz, 2013:5287; Jacobson, 2011), fue A.C. Ewing quien la estipuló claramente por primera vez al definir bueno como «el objeto idóneo de una actitud favorable hacia algo (...) ya sea desear, elegir, gustar, perseguir, aprobar y admirar»

(1947:149–152). Así, en la noción de «idoneidad» de la actitud está implicada una noción deóntica que no es equivalente a la obligación moral: no estamos moralmente obligados a tener una actitud positiva aunque reconozcamos que ella es una actitud idónea frente a ese objeto.

Al unir estas dos tesis, una CVA puede definirse de la siguiente manera: «x es valioso cuando puede ser objeto de una actitud o respuesta emocional idónea» (Steinfath, 2014:81).

El atractivo de CVA resulta claro al ubicarla en una taxonomía, aunque sea tosca, de las posibles concepciones respecto de la naturaleza de los valores. Siguiendo a Oddie (2015:60), hay dos clases de respuestas a la pregunta de qué son los valores: las respuestas idealistas y las realistas. Las primeras sostienen que no hay valores que sean independientes de nuestro deseo (o emociones, etc.) sino que son totalmente dependientes de nuestras actitudes. Las segundas sostienen que hay valores y hechos acerca de ellos que existen allende los deseos y que son quienes hacen las cosas deseables, dignas de desear, idóneas, etc. Así, la taxonomía se estructura entre un extremo idealista de que algo es envidiable porque lo envidiamos realmente y uno realista, de acuerdo con el cual hay cosas envidiables aunque no las envidiemos.

Dada su simpleza, esta taxonomía admite múltiples alternativas y especificaciones. Dos de ellas, como señala Jacobson (2011), son particularmente centrales e importantes para comprender las CVA. Tomemos como representante del extremo realista el realismo robusto de David Enoch, según quien los valores existen independientemente de las respuestas humanas a ellas, de forma análoga a cómo las propiedades primarias de la forma y el tamaño son lo que son independientemente de las reacciones humanas. Asimismo, tomemos como instancia del extremo idealista, el disposicionalismo defendido por Michael Smith, según el cual el valor es cualquier cosa que tiene el poder de causar ciertas reacciones en los seres humanos en condiciones estándar, de forma análoga a la que el color de un objeto es la combinación entre las propiedades microfísicas del objeto y su disposición a producir la visión en nuestras capacidades perceptivas. Así, mientras que, para el disposicionalismo, lo admirable es lo que produce admiración en los seres humanos, para el realista robusto la actitud de admiración es, cuando es correcta, una percepción adecuada de la calidad independiente de admirabilidad.

CVA se presenta como una combinación superadora de cada una de estas concepciones rivales. Por un lado, captura que debemos admirar ciertas cosas porque ellas son admirables pero, por el otro lado, defiende que no podemos entender lo admirable si no tomamos en cuenta nuestras respuestas y sensibilidad, es decir, nuestras capacidades para admirar. Puesto de otra manera, admite —junto con el realismo robusto— que puede haber casos donde la

admiración no rastree lo admirable y, al mismo tiempo, admite —con el disposicionalismo— que nuestra sensibilidad y disposiciones cumplen una parte esencial y no contingente en la comprensión de lo admirable.

En parte debido a esta síntesis superadora, las CVA presentan dos ventajas intrínsecas para comprender la naturaleza de los valores.

En primer lugar, no necesita de una mitificación de la naturaleza metafísica de los valores porque su carácter intrínsecamente prescriptivo —como lo llamaría John Mackie— está contenido en el propio valor y este es explicado en términos de razones para la idoneidad; después de todo, que algo sea valioso implica que debemos adoptar determinada actitud frente a él. Esto resulta una ventaja si se tiene en cuenta que George Moore y H. Prichard sostenían, por ejemplo, que los valores eran, al mismo tiempo, metafísicamente independientes y propiedades no naturales que motivaban intrínsecamente.

En segundo lugar, las CVA son metaéticamente neutrales tanto en la disputa entre cognitivistas y no cognitivistas como en la disputa entre realistas y antirealistas. Como señalan tanto Rabinowicz y Rønnow-Rasmussen (2004) como D'Arms y Jacobson (2000, 2006b), la segunda tesis de la CVA —la prioridad deóntica— admite interpretaciones de la idoneidad de la respuesta favorable frente al valor tanto de corte cognitivista como de corte no cognitivista:<sup>1</sup> si los juicios evaluativos son reducibles a las razones para adoptar actitudes favorables hacia un objeto, esas razones pueden ser tanto proposicionales como recomendaciones o expresiones de aprobación. Por ejemplo, las teorías de la sensibilidad como las de David Wiggins y John McDowell sostienen que un juicio evaluativo es adecuado cuando hay razones suficientes (cuando es idóneo) para adoptar una respuesta emocional determinada hacia el objeto en cuestión mientras que versiones sofisticadas del expresivismo (como el Expresivismo de Norma de Allan Gibbard y el casi-realismo de Simon Blackburn) sostienen que la idoneidad depende de la actitud de aceptación de una norma que haría apropiadas determinadas respuestas emocionales.

Como debería ser claro, el principal reto de CVA consiste en explicar la noción de Idoneidad, es decir, en explicar qué hace que determinadas razones para tener una actitud favorable hacia *x* sean apropiadas para realizar el juicio evaluativo específico sobre *x*. Más en detalle, esa noción de Idoneidad debería asumir que no cualquier razón para tener una actitud favorable (o desfavorable) hacia *x* es apropiada para realizar un juicio evaluativo sobre *x*. Por ejem-

---

1 Deonna y Teroni (2012:45–46) ofrecen una interpretación biologicista de esta idoneidad de las respuestas emocionales, siguiendo a Ruse y Wilson. Sauer (2011:126–130) reconstruye una interpretación evolucionista de la función de las emociones; por ejemplo, el miedo a las tarántulas podría ser apropiado para los seres humanos porque se corresponde con la función que nuestra forma biológica asignó al miedo: la de evitar el daño físico.

plo, yo puedo admirar a los editores de este libro y demostrárselo por escrito porque ello me permitirá tener una publicación más en mi currículum. Sin dudas, esa adulación autointeresada constituye una razón para esa actitud favorable hacia el editor; sin embargo, nadie diría que esa razón es idónea para juzgar la admirabilidad del editor. De esta manera, las CVA deben enfrentarse al problema de las Razones Erróneas, es decir, de que pueden existir un conjunto variopinto de razones para adoptar actitudes valorativas y que no todas ellas tienen que ver con las cuestiones valorativas que el análisis intenta resolver. Si este desafío no es enfrentado exitosamente por las CVA, su atractivo comienza a desmoronarse ya que perdería su capacidad para comprender el desacuerdo valorativo (Rabinowicz y Rønnow-Rasmussen, 2004); si el adulador autointeresado afirma la admirabilidad de los editores por razones estratégicas y alguien la niega señalando que sus trayectorias académicas son más que modestas, parece claro que no tienen un desacuerdo genuino sobre la admirabilidad. En la próxima sección, reconstruiremos la forma especialmente perjudicial que este problema adquiere para la CVA.

## **El Problema del Entrevero**

Como señalamos, la prioridad deóntica de las CVA les exige explicar el espacio crítico entre experimentar una emoción y que esa emoción sea idónea, es decir, entre que sintamos envidia y que esa emoción sea idónea ante una situación u objeto en particular.

Las CVA pretenden dar cuenta de este espacio crítico a través de lo que D'Arms y Jacobson (2000:729) denominaron la Tesis de Dependencia de la Respuesta (TDR). De acuerdo con ella, pensar que  $x$  tiene la propiedad  $P$  es pensar que es apropiado tener la actitud o emoción  $R$  como respuesta a  $x$ .<sup>2</sup>

Ahora bien, TDR corre dos riesgos evidentes, el de la Circularidad y el del Problema del Entrevero. El primero surge de que las razones que hacen idónea a determinada emoción frente a una situación deben ser especificadas de manera independiente a las propiedades evaluativas; es decir, que las razones para esa emoción no deben hacer referencia a la propiedad evaluativa que se intenta

---

2 El estatus de esta tesis es controvertido, aunque D'Arms y Jacobson son más bien ambiguos al respecto. Salmela distingue entre una versión semántica de la TDR (donde el objeto de análisis son los conceptos valorativos) y una versión metafísica (donde el objeto son las propiedades evaluativas). Al mismo tiempo, Tappolet y Rossi (2016:18) señalan que TDR solo es una tesis propia de CVA en la medida en que no se proponga como una mera elucidación conceptual (donde el concepto de admirable se conecte intrínsecamente al concepto de admiración) y, en cambio, se proponga como una tesis metafísica o semántica.



explicar. (Deonna y Teroni, 2012:46); así, uno no debería decir que es idóneo sentir envidia cuando se está frente a situaciones envidiables sino que, por el contrario, debería decir que es idóneo sentir envidia cuando se satisfacen un conjunto de razones y de contextos que hacen sentido que se sienta envidia.

Si bien el riesgo es claro, los propios defensores de la CVA parecen asumirlo directamente. Por ejemplo, Wiggins y McDowell insistían en que la circularidad implicada era, por un lado, informativa y, por el otro lado, inevitable. En primer lugar, la circularidad era benigna e informativa porque se produce al tomar «un desvío por los sentimientos» (Wiggins, 1987:187) que recuerda que los valores son incomprensibles si no están formulados para nuestra sensibilidad (tesis antropocéntrica); sería una práctica absolutamente extravagante preguntarnos qué significa lo admirable sin tomar en cuenta qué situaciones nos despierta admiración. En segundo lugar, la circularidad es inevitable porque creer que la admirabilidad puede comprenderse sin ninguna referencia al sentimiento en cuestión o que puedo comprender la admiración sin ninguna referencia a la admirabilidad implica recaer en los problemas del realismo robusto del que se quería escapar (Wiggins, 1987:195; McDowell, 1988:145).

El segundo riesgo es el Problema del Entrevero (PE), un tipo especial del problema de las razones incorrectas (Rabinowicz y Rønnow-Rasmussen, 2004: 402) y que se produce por la ambigüedad crucial con que una CVA se aproxima a términos como «idóneo», «apropiado», «garantizado», etc. Este problema consiste en que CVA, para ser exitosa, debe poder distinguir entre las numerosas razones o apoyos de una emoción únicamente aquellas razones que sean idóneas y pertinentes para establecer la propiedad valorativa en cuestión. Si las CVA no pueden realizar tal distinción, entonces podría «ofrecer respuestas sistemáticamente incorrectas a las preguntas sobre valores» (D'Arms y Jacobson, 2000:732), es decir, que las consideraciones relevantes para determinada respuesta emocional estarán entreveradas con las consideraciones irrelevantes.

Si bien los ejemplos que dan los autores son variados, uno de ellos es particularmente interesante. Dado que nuestro sentido del humor es dinámico y cambiante, las cosas con las que nos divertimos van cambiando a lo largo del tiempo. Dos payasos que se lanzan tortas de crema resultan hilarantes para niños pequeños pero un desperdicio de comida para los adultos; buscar palabras groseras en el diccionario ya no es tan gracioso como lo era apenas empezamos a leer. Sin embargo, no es solo el desarrollo madurativo lo que altera nuestro sentido del humor. También intervienen las normas sociales y la cultura de nuestra sociedad: no solo ya no es políticamente correcto realizar bromas sobre un supuesto desborde consumista en las mujeres sino que, ahora, las llamamos bromas machistas y, fundamentalmente para nuestros objetivos,

ya no nos entretenemos con ellas. De hecho, parecería que la broma machista no solo merece la censura moral sino que la emoción frente a ella tiende a ser, con buenas razones, la indignación.

Sin embargo, el Problema del Entrevero no surge por la actitud moral que deberíamos tomar hacia el chiste misógino; es claro que deberíamos condenar moralmente la diversión que despierta, ya sea en otros o en nosotros mismos. El problema es si esa condena (esa actitud de segundo orden) constituye una razón adecuada o idónea para juzgar la comicidad de la broma ya que, *prima facie*, parece totalmente razonable que una broma machista sea, no obstante, cómica. Ahora bien, supongamos que hay un desacuerdo sobre su comicidad, recordando que la posibilidad de dar cuenta de estos desacuerdos es una de las ventajas de las CVA. Por un lado, el cómico moralista sostiene que el machismo de la broma hace que la diversión no sea una respuesta idónea y que, por lo tanto, la broma no tenga la propiedad valorativa de ser cómica. Por otro lado, el cómico autonomista sostiene que aunque sea machista, la broma es *realmente* cómica porque la emoción de diversión es idónea frente a ciertas características formales del chiste y porque esas características son las únicas relevantes para juzgar su idoneidad. Así planteado, el desacuerdo es acerca de qué razones hacen que determinada emoción o actitud sea idónea para el objeto en cuestión; ambas razones (las morales y las cómico-formales) se refieren directamente al objeto en cuestión, son adecuadas (en algún sentido) para su evaluación y, no obstante, persiste un sentido en que esa aplicación de razones morales no parece tener que ver (*bear on*) con la comicidad de la broma.

Así, mientras que razones morales o prudenciales pueden aplicarse perfectamente para juzgar lo apropiado de determinadas emociones o actitudes, no parecen ser razones correctas para determinar cierta propiedad evaluativa; que nuestro amigo se enoje si envidiamos su auto no puede implicar que su auto no sea envidiable, que intentemos arrepentirnos menos de nuestras acciones no dice nada de si aquello de lo que no nos arrepentimos es o no realmente digno de arrepentimiento, etc. De esta forma, las CVA parecieran ser incapaces de distinguir entre razones que tienen que ver con cierta propiedad valorativa y las razones que no tienen que ver; unas se entrecruzan con las otras y las CVA corren el riesgo de emitir definiciones sistemáticamente incorrectas respecto de los valores. Las CVA necesitan, entonces, un criterio lo suficientemente rico como para cumplir con sus promesas.

## ¿Pueden ser cómicos los chistes misóginos?

El Problema del Entrevero amenaza la capacidad de las CVA para explicar la naturaleza de los valores ya que las expone a la amenaza de ser una estrategia frustrante que o bien necesita de una concepción previa y no justificada de qué es tener la propiedad evaluativa P o bien nos regresa al punto de partida a través de un desvío ya no por los sentimientos, sino por el ambiguo concepto de la idoneidad (Sauer, 2011:216). En este sentido, los intentos de los autores clásicos de las CVA (Wiggins, Blackburn, McDowell, Gibbard, entre otros) por dar criterios más precisos acerca de la idoneidad no parecen haber registrado la seriedad del problema.

Tomemos el intento de Gibbard por precisar el concepto de «emoción racionalmente garantizada (*warranted*)», que pareciera ser el más promisorio y el más consciente del problema. De acuerdo con él, las razones morales y estratégicas para tener ciertos sentimientos deben ser distinguidas de las que garantizan racionalmente tener tales sentimientos; las únicas razones que mostrarían la idoneidad de una emoción específica serían las últimas porque están apoyadas por la evidencia y no por consideraciones de utilidad o de moralidad. Así, aunque pueda ser racional *querer tener* determinada actitud (no divertirnos) frente al chiste misógino, la racionalidad de *tener* la actitud de no divertirnos estaría únicamente apoyada por consideraciones evidenciales; que yo refuerce los lazos grupales por no divertirme con ese chiste no es suficiente para mostrar la idoneidad de esa reacción emocional. Esa garantía evidencial debería ser buscada en la lógica interna de cada una de las emociones y, dado el Expresivismo de Norma que defiende Gibbard, en las propias aprobaciones (*endorsements*) del agente (1990:36-37). Esta distinción gibbardiana, no obstante, sigue siendo víctima del Problema del Entrevero porque pueden existir casos donde la garantía evidencial de una emoción se oponga a la aprobación del agente de esa emoción, esto es, donde el agente tenga razones evidenciales para tener una emoción x pero, no obstante, tengan más peso las consideraciones estratégicas para no tener esa emoción x y, por lo tanto, el razonamiento no sea útil para saber si la característica o situación en cuestión son o no idóneas para la emoción x (Salmela, 2014:149; D'Arms y Jacobson, 2000:740-747).<sup>3</sup>

---

<sup>3</sup> El ejemplo que D'Arms y Jacobson utilizan es el de un estudiante de filosofía que se siente avergonzado por sus mediocres talentos argumentativos (y esta emoción está garantizada evidencialmente) y que, al mismo tiempo, rechaza esa emoción porque sabe que obstaculizará cualquier progreso de su capacidad argumentativa. Si esto es así, el estudiante aprueba tanto su vergüenza como su no querer sentir vergüenza. Frente a este caso, la objeción es que la garantía evidencial de una emoción no es suficiente para determinar su idoneidad.

Frente a esta incapacidad de las CVA originales, en los últimos años surgieron numerosos y sofisticados proyectos que intentan ofrecer un criterio de distinción entre esas razones de idoneidad (Parfit, 2001; D'Arms y Jacobson, 2006b; Raz, 2011; Rabinowicz y Rønnow-Rasmussen, 2006), lo cual se debe, en parte, a que el problema de las razones incorrectas comenzó a ganar una atención fenomenal en la discusión ética y metaética. A continuación, solo reconstruiremos una de esas soluciones que consideramos que, al mismo tiempo que es mínimamente adecuada y simple, es sintomática del carácter casi aporético al que conduce tomarse en serio el PE.

Esta solución fue propuesta por Danielsson y Olson (2007). Su tesis central es que debemos distinguir entre las razones que se refieren al *contenido* de una actitud frente a un determinado objeto y las razones que se refieren a la *actitud* frente a determinado objeto (*holding-reasons*); mientras que las primeras se aplican para determinar la corrección o la verdad de la actitud frente al objeto, las segundas pueden aplicarse a la conveniencia o a la moralidad de tener esa actitud. Esta distinción permitiría evitar inferencias erróneas como las que constituyen el Problema del Entrevero; mientras que a partir de una razón de contenido puede deducirse una razón de actitud (y, así, constituirse como una razón correcta), no se da el caso contrario. Volviendo al caso del humor misógino, el chiste puede ser gracioso por razones de contenido y, no obstante, el agente puede tener razones de actitud para no divertirse con él; más importante para nuestros objetivos, tener esas razones de actitud no sería suficiente para establecer que el chiste no sea divertido. Por lo tanto, las razones incorrectas serían aquellas razones para determinada actitud que no se siguen de la existencia de razones de contenido.

La solución, sin embargo, tiene un carácter meramente formal que pone en duda su carácter informativo; el criterio que carga con el peso de la distinción solo es especificado utilizando las propiedades descriptivas de los objetos que dan razones de contenido para las actitudes evaluativas. Así, según Louise (2009:353), la solución de Danielsson y Olson produciría la siguiente definición: «A es gracioso si y sólo si A tiene las propiedades que hacen que la diversión hacia x sea idónea, esto es, que den razones de contenido para divertirnos hacia x».

Si esto es todo lo que esta solución puede ofrecer, aún no contamos con ninguna idea de cómo seleccionar cuáles son las propiedades descriptivas de los objetos que ofrecen razones de contenido para la idoneidad de las respuestas emocionales. Por el contrario, tan solo conseguimos una recharacterización de la idoneidad formal de las razones para que algo sea cómico, que no se derivan de razones de actitud.

Este carácter poco informativo podría motivar dos caminos distintos. Podríamos creer que el problema es local y que debemos intentar alcanzar otros criterios de distinción para la idoneidad de determinada emoción frente a cierta propiedad evaluativa o bien podríamos tratar de comprender las causas de este carácter poco informativo. Creemos que este segundo camino es el más prolífico y que nos conduce a investigar si los límites autoimpuestos por las CVA terminan causando el Problema del Entrevero. En particular, la cuestión será la de investigar si la tesis de la prioridad deontica es satisficible, como supuestamente lo es, con cualquier posición metaética, ya sea realista o idealista.

Cuando intentamos trascender este carácter poco informativo de la solución de Danielsson y Olson, nos hallamos de frente al centro neurálgico de las CVA, esto es, investigar las propiedades evaluativas únicamente en la corrección de una respuesta idónea, la cual no depende de las propiedades del objeto en un sentido interesante. Este centro neurálgico siempre estará amenazado por la posibilidad de que el objeto en cuestión tenga propiedades evaluativas que, por distintas razones, no sean captadas por el agente; el chiste puede ser cómico y el agente no entretenerse, el hogar de mi amigo envidiable y yo no sentir envidia, mi indiferencia hacia los otros vergonzosa y, no obstante, ser indiferente a la mirada del otro, etc. Sin concentrarse en esas propiedades del objeto, las soluciones al Problema del Entrevero postuladas dentro de las CVA y su análisis parecen destinadas a ese carácter poco informativo. Como señalan Deonna y Teroni:

la especificación de las razones relevantes no apela a las propiedades normativas de la respuesta sino a las propiedades evaluativas del objeto que se supone debemos analizar. Después de todo, es porque los objetos son admirables que la admiración hacia ellos es una respuesta correcta. (...) las emociones razonables o apropiadas son evaluaciones que representan justificadamente a sus objetos como teniendo las propiedades evaluativas relevantes. (2012:48)

Vale la pena aclarar que la asunción de esta tarea no significa directamente la defensa de algún tipo de posición realista. Recordemos que las propias CVA sostienen la intuición de que pueden existir cosas valiosas que, no obstante, no sean consideradas valiosas por los agentes. En este sentido, pareciera que las CVA solo se concentran en un aspecto de las adscripciones evaluativas (las respuestas de los agentes), condenándose a no poder explicar la distinción entre que algo sea valioso y que el agente no tenga la respuesta idónea hacia él.

El caso de los chistes misóginos ejemplifica ese problema. Las propiedades de este objeto (su misoginia) son asumidas como datos que no afectan a otras

propiedades evaluativas (su comicidad) y la respuesta del CVA al Problema del Entrevero debería explicar por qué cada una de estas propiedades genera una respuesta emocional diferente y autónoma. Sin embargo, si esas propiedades no son analizadas más profundamente, el Problema del Entrevero se agiganta ya que implicaría que nunca hay una razón para afirmar que una broma racista no sea cómica por ser racista, es decir, que ninguna consideración moral es relevante para juzgar la propiedad evaluativa de su comicidad.

Sin embargo, como mostraron Jordan y Patridge (2012), esto no es siempre así: pueden existir casos donde las razones morales sean idóneas para juzgar ciertas propiedades evaluativas. Como intentan mostrar, esto no implica adoptar el moralismo cómico, según el cual hay una inferencia directa desde la incorrección moral de la respuesta asociada a la broma (divertirse) al juicio de que la broma misógina no sea cómica.

De acuerdo con los autores, las consideraciones morales pueden silenciar (es decir, pueden ser una condición inhibidora de razones contributivas) las razones por las cuales el chiste misógino sea cómico si y solo si están presentes rasgos moralmente relevantes del contexto donde se realiza la broma. Si las razones por las que un agente se divierte con una broma están silenciadas por rasgos del contexto, entonces la diversión no es una respuesta idónea hacia ella y, por lo tanto, la broma no es cómica (en ese contexto). Como señalan:

las consideraciones morales cumplen el rol de inhibir cualquier razón que el agente podría considerar para estar entretenido, y esto incluye razones relevantes para la idoneidad. De esta manera, el agente no cuenta con ninguna razón para estar entretenido y como resultado el entretenimiento no es idóneo para su objeto (...) Algunas consideraciones morales que se aplican a si debe sentirse una emoción pueden ser, *pace* D'Arms y Jacobson, relevantes a la cuestión de la idoneidad. (Jordan y Patridge, 2012:93)

Por ejemplo, si el presidente Alberto Fernández pronuncia un discurso en el Congreso de la Nación y cuenta un chiste misógino, es claro que una lectura moral del contexto silencia las razones por las que el chiste pudiera ser cómico. En este caso no es solo que la respuesta emocional de divertirse no sería adecuada moralmente, sino que dejarían de estar activas las razones para considerar a esa broma como cómica. Resulta importante enfatizar que el carácter misógino del chiste no es la condición inhibidora (ya que eso sería caer en el moralismo cómico), sino los rasgos morales del contexto; dado esto, incluso si el presidente Alberto Fernández empezase su alocución en el Congreso preguntando por qué se suicidó un libro de matemáticas, también podríamos decir que el chiste no es cómico.

## Conclusión

Imaginemos una historia plausible del auge y caída de un chiste misógino. En el principio de los tiempos, era hilarante para todos y todas y aquellas personas que no lo consideraban cómico (en su mayoría, mujeres) eran consideradas amargadas. Luego, esas críticas empezaron a oírse más y los cómicos machistas respondían diciendo que «solo era una broma», que no había motivo para ofenderse. Algunas décadas más adelante todavía, el chiste comenzó a ser respondido con un «eso no es gracioso», es decir, que no solo condenaba moralmente la broma sino que cuestionaba que siquiera fuera cómico, la emoción de entretenimiento fuese idónea; en esta etapa, el chiste misógino pasa a ser un estereotipo machista, un síntoma para detectar individuos misóginos, etc. Finalmente, el chiste, dado que ya no cumple su función intrínseca, desaparece de la práctica social del humor. En esta historia hipotética pero plausible, hay una dialéctica entre el contexto donde se enmarca el objeto evaluativo y la emoción o actitud que es idónea, es decir, que la comicidad de la broma misógina no solo depende de las propiedades formales del objeto sino del contexto donde se cuenta.

Como mencionamos en la última sección, el centro neurálgico del problema de CVA es centrarse únicamente en la respuesta emocional o actitudinal frente a la broma racista y poner en un plano secundario al objeto evaluativo para definir su idoneidad. Este centro, que aleja a las CVA del realismo, también las distancia de una lectura integral del contexto donde se realizan las predicaciones valorativas y esto hace que se pierdan muchos elementos importantes para definir la idoneidad de una respuesta evaluativa. En la historia hipotética que presentamos, al cambiar ese contexto práctico integral, al modificarse el significado, los elementos salientes (por ejemplo, concentrarse ya no en quienes se divierten sino en quienes son el objeto material de la broma), también se modifica el objeto: ese chiste ya no solo no es gracioso, sino que tampoco es un chiste.

Esta incapacidad del centro neurálgico de las CVA es instructiva, al menos a nuestros ojos, para la construcción de una concepción sobre la naturaleza de los valores. Lo que nuestra historia plausible de la caída de los chistes misóginos muestra es que la tesis antropocéntrica respecto de los valores debe ser prioritaria respecto de la tesis de la prioridad deóntica, es decir, que los criterios por los cuales ciertas actitudes son idóneas para determinado objeto no pueden establecerse si no se toma en cuenta el contexto práctico y valorativo en el que se producen. Así, lo que determina si la respuesta emocional de divertirse es idónea o no para un chiste no son ni las propiedades formales del objeto ni la capacidad de derivar (pero no ser derivadas por) razones

de actitud, sino el conjunto de prácticas, valores y presiones sociales de una determinada comunidad práctica. Sin dudas, esto nos coloca más cerca del extremo idealista que del extremo realista. Creemos que el fracaso de las CVA para posicionarse como la solución superadora de las disputas entre el idealismo y el realismo exige, precisamente, una posición como la adoptada aquí: que no podemos escapar de la pregunta del *Eutifrón* y que la tarea debe ser la de argumentar directamente a favor o en contra de alguna variante realista o alguna idealista.

## Referencias bibliográficas

- D'Arms, J. y Jacobson, D. (2000). Sentiment and Value. *Ethics*, 110(4), 722–748.
- D'Arms, J. y Jacobson, D. (2006a). Anthropocentric constraints on human value. En Schafer-Landau, R. (Ed.), *Oxford studies in metaethics* (pp. 99–126). Vol. 1. Clarendon Press.
- D'Arms, J. y Jacobson, D. (2006b). Sensibility Theory and Projectivism. En Copp, D. (Ed.), *The Oxford Handbook of Ethical Theory* (pp. 186–218). Oxford University Press.
- Danielsson, S. y Olson, J. (2007). Brentano and buck-passers. *Mind*, 116(443), 511–522.
- Deonna, J. y Teroni, F. (2012). *The emotions: A philosophical introduction*. Routledge.
- Ewing, A.C. (1947). *The Definition of Good*. Routledge & Kegan Paul.
- Fronzizi, R. (1958). *¿Qué son los valores? Introducción a la Axiología*. Fondo de Cultura Económica.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Harvard University Press.
- Jacobson, D. (2011). Fitting Attitude Theories of Value. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/spr2011/entries/fitting-attitude-theories/>
- Jordan, A. y Patridge, S. (2012). Against the Moralistic Fallacy: A Modest Defense of a Modest Sentimentalism about Humor. *Ethical Theory and Moral Practice*, 15(1), 83–94.
- Louise, J. (2009). Correct responses and the priority of the normative. *Ethical Theory and Moral Practice*, 12(4), 345–364.
- McDowell, J. (1998). *Mind, Value, and Reality*. Harvard University Press.
- Oddie, G. (2015). Values and Desires. En Hirose, I. y Olson, J. (Eds.), *The Oxford Handbook of Value Theory* (pp. 60–79). Oxford University Press.
- Parfit, D. (2001). Rationality and Reasons. En Egonsson, D.; Petersson, B.; Josefsson, J. y Rønnow-Rasmussen, T. (Eds.), *Exploring Practical Philosophy: From Action to Value* (pp. 17–39). Ashgate.
- Rabinowicz, W. y Rønnow-Rasmussen, T. (2004). The Strike of the Demon: On Fitting Pro-attitudes. *Ethics*, 114(3), 391–423.
- Rabinowicz, W. y Rønnow-Rasmussen, T. (2006). Buck-passing and the right kind of reasons. *Philosophical Quarterly*, 56(222), 114–120.
- Rabinowicz, W. (2013). Fitting-Attitude Account of Value. En LaFollette, H. (Ed.), *The International Encyclopedia of Ethics* (pp. 1–12). Blackwell Publishing.
- Raz, J. (2011). *From Normativity to Responsibility*. Oxford University Press.



- Salmela, M. (2014). *True Emotions*. John Benjamins Publishing.
- Sauer, H. (2011). The Appropriateness of Emotions: Moral Judgment, Moral Emotions, and the Conflation Problem. *Ethical Perspectives*, 18(1), 107–140.
- Steinfath, H. (2014). Emociones, Valores y Moral. *Universitas Philosophica*, 31(63), 71–96.
- Tappolet, Ch. y Rossi, M. (2016). What is value? Where does it come from? A philosophical perspective. En Brosch, T. y Sander, D. (Ed.), *Handbook of Value*. Oxford University Press.
- Wiggins, D. (1987). *Needs, Values, Truth: Essays in the Philosophy of Value*. Blackwell.

## 21. Notas sobre el debate Kohlberg vs. Gilligan: de la diferencia de género a los desacuerdos metaéticos

*Adriana Vercellone*

Podríamos decir que las personas aprendemos a pensar moralmente desde la infancia, gracias a la interacción con el entorno, la familia y la sociedad. Que paralelamente al crecimiento biológico, desarrollamos la capacidad de razonar acerca de nuestros derechos y deberes, y a ponderarlos cuando alguna situación cuestiona las intuiciones y mandatos que nos rigen. Que el cumplimiento de las reglas sociales y la inquietud por habitar un mundo justo son el límite que contribuye a que reflexionemos acerca de todo lo anterior. Que, pese a las diferencias culturales, el desarrollo cognitivo y moral que alcanzamos es *más o menos* similar.

Tal descripción aúna trozos de los diferentes argumentos que han moldeado la concepción tradicional acerca del desarrollo moral de las personas. Al menos hasta hace algunos años, cuando algunas preguntas éticas fundamentales —*¿qué nos motiva a actuar?, ¿cómo reconocemos aquello que es correcto?*— tomaron un tono diferente en las ideas de Carol Gilligan.

Al indagar en los estudios de las principales teorías acerca del razonamiento moral, Gilligan encuentra que todas arriban a una conclusión arbitraria: las niñas y mujeres siempre se encuentran en una etapa más baja de desarrollo que los varones. Para mostrar que se trata de una afirmación androcéntrica insostenible, nos ofrece una novedosa propuesta que reconoce la experiencia femenina en la definición de la identidad y del juicio moral.

El principal blanco de tiro de su teoría son los casos de estudio y conclusiones de Lawrence Kohlberg. Pero la revolución que inició *In a different voice* (1982) trascendió dicho espectro, al punto de poner en tela de discusión otros asuntos, como algunos presupuestos metaéticos de la filosofía tradicional, o conceptos básicos de la psicología moral y del lenguaje. Así, la discusión suscitada entre Gilligan y Kohlberg es sobre género y desarrollo moral; pero también sobre otros temas que interesaron siempre a la filosofía. En ese sentido, pocos debates revelan la interseccionalidad del iniciado por Gilligan. Y aunque existe toda una tradición de teoría feminista que le precede, a ella debemos el haber introducido la preocupación por las relaciones y el cuidado en la filosofía moral.

El objetivo de estas páginas es repasar algunos puntos salientes del debate. Para ello primero se reconstruyen los dos modelos de razonamiento moral mencionados: el de Kohlberg y el de Gilligan. Aunque la referencia a ambas

teorías es breve, se procura resaltar el rol que conceden a la diferencia sexual para el desarrollo moral y psicológico de las personas. En este texto se sugiere que dicho rasgo constituye solo uno de los puntos de desacuerdo entre ambas posiciones, pues existe también una lista de presupuestos metaéticos y normativos que les brindan apoyo y resulta interesante explorar. Sobre ese tema avanza la segunda parte del capítulo, que despeja cinco interrogantes acerca del razonamiento moral, todos relevantes para un abordaje filosófico completo.

## **El desarrollo moral: de Kohlberg a Gilligan**

Las preguntas sobre el desarrollo moral ocupan una parte destacada en el campo de la ética. Estas indagan acerca de los cambios y progresos que atraviesan las personas desde su infancia y el modo en que evoluciona su entendimiento acerca de la moralidad. Las teorías sobre el tema no buscan responder «cómo formulamos una teoría moral» ni «qué procesos nos permiten reconocer principios morales válidos», sino «qué procesos psicológicos e identitarios moldean nuestra moralidad» (aunque, como se verá, todas estas preguntas guardan cierta vinculación).

Una de las respuestas más completas sobre el tema la ofreció Lawrence Kohlberg, al vincular el desarrollo psicológico y cognitivo de una persona con su desarrollo moral. Su propuesta se apoya, por un lado, en estudios cognitivo-evolutivos acerca de la socialización y el aprendizaje (1). Y, por otro lado, en importantes antecedentes sobre el desarrollo de la identidad (2).

La etiqueta de cognitivo-evolutiva (1) responde a una concepción interaccionista de los procesos mentales de aprendizaje de una persona. Entre sus principales exponentes destaca, sin dudas, J. Piaget para quien la interacción con el entorno es vital en el desarrollo de la capacidad de estructurar e internalizar normas. De acuerdo con esto, los niños y las niñas tienen estructuras mentales que van moldeando en función del entorno y su búsqueda por respuestas correctas; o, mejor dicho, en su búsqueda por comprender y adaptarse en el mundo. Este proceso continuo de estímulo y respuesta con el medio permite desarrollar construcciones mentales que, desde la infancia, se suceden de forma secuencial, invariable y estructurada (Kohlberg, 1992:33-34). Según Kohlberg, el desarrollo moral tiene lugar a la par de estos aspectos no morales del desarrollo cognitivo. Por ello, la capacidad para razonar moralmente y el movimiento de un estadio de madurez moral a otro es también el resultado de la interacción con el entorno y el tejido social.

El segundo punto de apoyo de la teoría (2) remonta a paradigmas psicológicos que identifican algunos sucesos que las personas atraviesan en las etapas

tempranas de la infancia y adolescencia, con los conflictos que desencadenan avances en el crecimiento y la percepción de la propia estima (Gilligan, 1993:17 y ss.). El desarrollo de la identidad se vincularía con estas crisis que ponen en tensión la independencia y la autonomía y, correlativamente, constituyen una oportunidad para desplegar vías de resolución que gradualmente se van complejizando y sofisticando.

Así, el desarrollo del razonamiento moral no puede desentenderse de los procesos de construcción de la identidad y de las capacidades para el pensamiento lógico abstracto.<sup>1</sup> Para descubrirlo, Kohlberg analiza diferentes entrevistas realizadas a menores y adolescentes, agrupando sus respuestas en función de categorías teóricas. Estos «tipos ideales» revelan las características de la secuencia de niveles de desarrollo moral propias para cada edad. Concretamente, se plantean dilemas hipotéticos a niños y niñas de distintas edades, cuyas reacciones permiten conocer y descubrir las estructuras de razonamiento y lógica interna de sus respuestas (Kohlberg, 1992:33-34).

En concreto, distingue tres niveles, que se subdividen en dos, y configuran en total las seis etapas de progreso moral que las personas pueden desarrollar. Las primeras dos etapas se corresponden con un nivel de desarrollo *premoral*, caracterizado por un entendimiento individualista o egocéntrico del mundo y los problemas. En esta primera instancia, el propio interés es aquello que guía las acciones individuales, siendo el castigo y el intercambio instrumental las principales razones que motivan a las personas a actuar. Algunas crisis comunes en la infancia abren paso a una segunda etapa —*convencional*— en la que existe una visión societal o comunal de los problemas, guiada por la búsqueda y mantenimiento de buenas relaciones y de la autoridad. Sin embargo, en dicho estadio aún no se ha desarrollado una acabada noción de los derechos y los valores como entes colectivamente contruidos, facultad que se desarrolla solo en las quinta y sexta etapas del tercer nivel, llamado *posconvencional*.

Resume los seis tipos de estadio moral evolutivo del siguiente modo (Kohlberg, 1992:35):

---

1 Aunque existen paralelismos entre estos procesos, no siempre es perfecto. Según Kohlberg, alguien en una etapa de desarrollo cognitivo X pudiera encontrarse en una etapa inferior de desarrollo moral.

Tabla 1

NIVEL A. PREMORAL
Estadio 1. Orientación al castigo y obediencia.
Estadio 2. Ingenuo hedonismo instrumental.
NIVEL B. MORALIDAD DE CONFORMIDAD CON EL ROL CONVENCIONAL
Estadio 3. Moralidad de mantenimiento de buenas relaciones, aprobación por parte de otros.
Estadio 4. Moralidad del mantenimiento de la autoridad.
NIVEL C. MORALIDAD DE PRINCIPIOS MORALES AUTOACEPTADOS
Estadio 5. Moralidad de compromiso, de derechos individuales y de la ley democráticamente aceptada.
Estadio 6. Moralidad de principios individuales de conciencia.

Lo anterior revela una estructura invariable y jerarquizada en la que el desarrollo de una aptitud moral desplaza a la anterior (Kohlberg, 1992:35). Dicha secuencia se supone universal, pues cada quien debiera atravesarla sin saltarse ni retrotraer ningún estadio y, finalmente, alcanzar el último: un sexto nivel superior de desarrollo moral que, de hecho, no todas las personas logran.

Al evaluar los resultados de las entrevistas, Carol Gilligan nota que las conclusiones de Kohlberg silencian aquellas respuestas que revelaron valores morales y atributos propiamente «femeninos». Ya que, al no adecuarse a las réplicas de los tipos ideales, estas fueron señaladas como defectuosas o en un nivel inferior. Para entender mejor esta observación, es necesario prestar atención no solo a la diferencia entre las entrevistas, sino también a otros desacuerdos que surgen en el diálogo con los dilemas.

Consideremos primero el dilema hipotético más recurrido en el estudio:

*Heinz se debate entre robar o no una droga que su esposa —gravemente enferma— necesita para salvar su vida, pero no puede costear. ¿Sería incorrecto que Heinz robe la medicina?*

Gilligan comienza por destacar que jóvenes de diferentes edades y género —entre quienes presta especial atención a Jake y Amy— interpretan disímelmente el problema planteado. Allí donde Jake encuentra un conflicto entre dos valores básicos (propiedad y vida), Amy percibe comprometidas las relaciones que vinculan a las personas en cuestión: la relación de Heinz con el farmacéutico y, fundamentalmente, la de Heinz con su esposa, que es de preservación y afecto (Gilligan, 1993:26–28).

Asimismo, Gilligan nota que la causa del dilema se percibe diferente en ambas entrevistas. Mientras Jake cree que existe un problema por la falta de acceso al medicamento, hecho que puede comprometer el cumplimiento de un deber (pues si Heinz decide robar entonces incumple un deber para con el

farmacéutico), Amy encuentra que algunas de las personas implicadas pudieran haber realizado más para contribuir con el bienestar ajeno (el farmacéutico, por ejemplo, podría bajar el precio del producto) (Gilligan, 1993:27-28). Y lo mismo sucede con la posible solución al conflicto, pues Jake relata una tensión entre deberes individuales que puede ser resuelta lógicamente, mientras que Amy se preocupa por preservar las relaciones en juego y evitar el mayor daño a todas las personas implicadas. Inclusive, Gilligan destaca que la actitud ante la propia entrevista es diferente en uno y otra: Jake interpreta las preguntas de forma rigurosa, en búsqueda de respuestas positivas o negativas; mientras que Amy descubre un diálogo acerca de «*cuál es el mejor modo de comportarse en una situación determinada*».

Todo esto, evidencia la mirada jerarquizada que Jake tiene acerca de los problemas y las relaciones de poder y, por el contrario, una visión relacional de la situación por parte de las mujeres, que incluye a la comunicación como eje fundamental en su búsqueda de soluciones. En resumen, las mujeres no evalúan el dilema como asuntos que desafían a personas aisladas ni buscan respuestas en algún sistema de reglas dado; más bien, consideran a personas interconectadas cuyas relaciones es necesario preservar (Gilligan, 1993:28-32). Así lo explicita al evaluar otras dos entrevistas: «pero mientras Jeffrey piensa qué va primero, Karen se concentra en quién es excluido» (1993:33).

Las diferencias en el desarrollo moral que presentan hombres y mujeres recorren todo el texto de Gilligan y descubren la explícita exclusión que Kohlberg hace de la opinión femenina en los problemas planteados. Tanto las entrevistas de Jake y Jeffrey como la respuesta masculina promedio sirvieron para construir los niveles de desarrollo que son «normales» o aceptables para Kohlberg; mientras que, las de las mujeres, no alcanzaron el estadio mayor de madurez moral dadas sus diferentes reacciones ante los dilemas (1993:27).

Así, sobre la base de datos de entrevistas y problemas hipotéticos similares, Gilligan concluye tres niveles de reflexión diferentes a los de Kohlberg; asimismo, distingue dos factores de transición que influyen en el traspaso de un nivel a otro. Esquema que, con claridad, pone en evidencia la influencia de la diferencia sexual en el desarrollo moral y psicológico.

El nivel inicial es de mera supervivencia, caracterizado por el cuidado y aseguramiento de sí misma. Una etapa de transición en la que se cuestiona el pensamiento egoísta y se considera el valor de la bondad, abre paso a un segundo nivel de desarrollo moral. En dicho estadio, deja de ser prioridad el cuidado de la propia persona, y la atención se pone sobre los demás. Tal exclusión lleva a cuestionar las razones por las que la opinión e intereses ajenos debieran ser tan relevantes, transición que conlleva a una tercera etapa, centrada en la dinámica de las relaciones. Este último estadio resuelve el dilema

responsabilidad–egoísmo característico de las instancias anteriores, a través de un genuino entendimiento de la interconexión entre sí y las demás personas. Finalmente, esto da lugar al desarrollo de una «ética el cuidado» (Gilligan, 1993:74–75).

Para resumirlo, considérese el siguiente esquema:

Tabla 2

NIVEL A. SUPERVIVENCIA: Egoísmo y autointerés.
Factor de transición 1. Cuestionamiento del pensamiento egoísta y reconocimiento de la responsabilidad hacia los demás.
NIVEL B. RESPONSABILIDAD HACIA LOS/LAS DEMÁS: Adecuación del propio comportamiento a expectativas sociales, supresión de las propias creencias e intereses.
Factor de transición 2. Desequilibrio por la exclusión de la propia persona.
NIVEL C. INTERDEPENDENCIA CON LOS DEMÁS Y CUIDADO: Preocupación por minimizar el daño a los/las demás, sin despreciar la responsabilidad hacia sí misma.

Las diferencias entre ambas tesis pueden rastrearse también en sus bases teóricas. En dicho sentido, Gilligan destaca el modo en que la teoría psicológica y del desarrollo moral han estudiado las experiencias familiares, así como las crisis que suceden en el crecimiento personal (1993:4–16). Según la filósofa, el foco se ha puesto siempre en dos sucesos: la cercanía y la separación con la madre —u otra figura de apego—. Según las teorías tradicionales (androcéntricas), al no atravesar experiencias trascendentales para su desarrollo —como el desapego de la madre o el complejo de Edipo— las mujeres se ven privadas también de una instancia de consolidación de su identidad.

Sin embargo, Gilligan interpreta ese hecho desde una óptica diferente: allí donde los hombres experimentaron una crisis de separación con el hogar que les abrió paso al ámbito público, las mujeres encontraron lo contrario, una privación —de oportunidades— en lo público y un apego familiar mayor. Gilligan nota que los varones entrevistados vinculan sus procesos de logros personales con períodos de separación respecto de terceros, instancias percibidas como condición de libertad y autodesarrollo (1993:152–167). Mientras que para las mujeres la separación y la cercanía representaron imágenes de indefensión y falta de poder, que derivan en la contradicción integridad–cuidado. En definitiva, experiencias que sirven a la maduración personal, construcción de la identidad y moldean la moralidad se viven y perciben diferenciadamente según el género.

Esto permite a Gilligan exponer el modo en que las teorías psicológicas y cognitivas más difundidas hallan en la diferencia biológica un déficit para el

desarrollo de las mujeres. El problema, en definitiva, ha sido que el orden secuencial de desarrollo elaborado por la tradición teórica más difundida se corresponde con la percepción masculina acerca del desarrollo. Construido sobre rupturas y separaciones en la primera infancia y la adolescencia, ha sido un modelo que se contraponen con la experiencia femenina, construida sobre conexiones y relaciones continuas (Gilligan, 1993:39, 45-48).

Todo lo anterior permite a Gilligan elaborar una aguda crítica a los estudios de Kohlberg, así como construir una teoría acerca del desarrollo que rescata la experiencia femenina en la definición de la identidad y el juicio moral.

### Dos modelos éticos: de la justicia al cuidado

La historia familiar de las mujeres entrevistadas reveló algunas características particulares. Principalmente, una gran tensión entre el compromiso consigo misma —las propias preferencias— y el compromiso hacia los demás. Llegado cierto momento de la vida, las mujeres, usualmente formadas para el sacrificio y las relaciones privadas, introdujeron entre sus preocupaciones a *los derechos* (esto es, la preocupación liberal por el desarrollo personal y la no interferencia). Según Gilligan, este diálogo tuvo el efecto de concientizar respecto de la abnegación que la sociedad tradicionalmente les ha exigido (1993:129) o el modo en que la virtud femenina se ha identificado con el autosacrificio (132). Asimismo, sacó a la luz la falta de voz que las mujeres tienen en el espacio público (149) e incluyó en la ecuación la cuestión de la responsabilidad por las relaciones sociales (129). No es casual que Gilligan encuentre que las grandes crisis de transición se desencadenan con la maternidad, la crianza o elecciones radicales como la del aborto (108 y ss.); pues son momentos que ponen en conflicto las esferas de la responsabilidad hacia los demás con los derechos individuales; del cuidado y la intimidad, con el desarrollo personal y los reclamos de equidad.

Sin embargo, lo anterior no implicó que las mujeres abandonen su preocupación por las relaciones humanas y la responsabilidad. De hecho, la disputa entre el discurso de los derechos y la justicia con la ética del autosacrificio es lo que les permite percibirse como interdependientes de sus relaciones e, inclusive, reconocer que en la comprensión de dicha conexión reside una gran fortaleza (Gilligan, 1993:147-149). Así, en el último estadio de desarrollo emerge una concepción de moralidad que involucra todas las aristas mencionadas de forma integral. Es la llamada «ética del cuidado», que exige actuar con responsabilidad hacia sí y las demás personas (147-149); centrada en hacer del mundo uno más seguro y encontrar respuestas colectivas no violentas ante los problemas.



Las éticas del cuidado contrastan de manera evidente con la concepción tradicional acerca de los derechos, que Kohlberg identifica con el estadio 6 de desarrollo moral. Esta última, erigida sobre el ideal de la justicia, concibe a los problemas morales como dilemas de deberes individuales que se les presentan a agentes aislados, los cuales pueden ser resueltos apelando a la lógica y las categorías formales de las normas. Por el contrario, la ética del cuidado posiciona a las personas en relación con el mundo, contextualizando la situación y evaluando las relaciones y conexiones que se hallan en juego (Gilligan, 1993:35-37). Allí donde una reconoce a los derechos como el límite a las elecciones individuales, la otra encuentra soluciones en el diálogo y las categorías de la responsabilidad.

En este punto resalta la crítica de Gilligan al modelo tradicional acerca del desarrollo moral humano. Cuando los conflictos se perciben de forma jerarquizada y las respuestas se buscan en un sistema lógico que balancea deberes y justicia, se denigra la preocupación por la responsabilidad y el cuidado de las relaciones; es decir, aquello que caracteriza al razonamiento femenino. Las respuestas de algunos entrevistados se consideran en «un nivel más alto (...) sólo si se supone, como lo hace Kohlberg, que una ética de los principios es superior a una ética que subraya la intimidad, la atención a la persona y las relaciones personales» (Rachels, 2006:250).

### **Diferencia de género, presupuestos metaéticos y normativos**

Como se expresó, las teorías sobre el razonamiento y el desarrollo moral indagan acerca de los procesos que, desde la infancia, moldean y forman la capacidad de comprender y argumentar acerca de la moralidad. Es decir, estudian el desarrollo de los procesos cognitivos y psicológicos que permiten a las personas pensar moralmente, sea en torno al significado de los juicios morales como a su justificación.

Las respuestas sobre el tema son diferentes a las elaboradas en el plano de la *metaética* —es decir, la indagación por el significado de los juicios y conceptos morales, así como por sus procedimientos de validación y verificación— y de la *ética normativa* —que interroga por la justificación de los juicios de valor, procurando elaborar principios que sirvan de guía para la acción humana—. No obstante, es inevitable que dichos niveles de reflexión se solapen. En ese sentido, tanto el argumento de Kohlberg como el de Gilligan se construyen sobre presupuestos metaéticos y normativos fundacionales.

A continuación, consideramos cinco preguntas que procuran dilucidar dichas conexiones: (a) ¿qué tipo de verdades postulan los juicios morales?, (b)

¿cómo se formulan las verdades morales?, (c) ¿qué significan los términos de la moral?, (d) ¿qué tipo de ejercicio mental supone el razonamiento moral? y (e) ¿el desarrollo de aptitudes morales nos conduce hacia verdades morales?

### **¿Qué tipo de verdades postulan los juicios morales?**

En torno al tipo de verdades que encierran los juicios morales, Kohlberg reconoce un inicial compromiso con la universalidad (1992:281 y ss.). Tal cosa sucede tanto en lo que refiere a las estructuras de razonamiento, como también en relación con los principios que se consideran aceptables (es decir, los asuntos normativos). En otras palabras, al describir las etapas de desarrollo moral humano, Kohlberg asume que estas son universales y constituyen un camino hacia el reconocimiento de un discurso moral uniforme.

La discusión en este punto se suscita con el llamado «relativismo moral y cultural» que niega la existencia de un criterio independiente de corrección o incorrección (Rachels, 2006:39-40). Según el relativismo, los juicios morales no son otra cosa que verdades dependientes de convenciones humanas o relativas a cada sociedad y tiempo histórico.

Al discutirlo, Kohlberg reconoce que existen evidentes diferencias culturales entre sociedades; aunque asume que estas se suscitan en torno a conductas y costumbres, más no respecto de valores básicos. Por ejemplo, sugiere que algunas prácticas sexuales —como la monogamia o la poligamia— pueden depender de la cultura y los hábitos, si bien todas son compatibles con la defensa de un principio básico de dignidad humana, confianza y compromiso en las relaciones sexuales (1992:282). Así, Kohlberg defiende la prioridad de valores humanos básicos, umbrales que deben ser respetados a la hora de decidir cómo actuar.

Semejante aseveración a nivel normativo supone también la existencia de estructuras de razonamiento que son las que permiten reconocer y formular principios morales. Dicha estructura incluye, según Kohlberg, la aprehensión de habilidades para el pensamiento lógico y abstracto, la capacidad imaginativa y la de elaborar categorías conceptuales *significativas* (Thomas, 1995:622). Estas forman parte de una secuencia de desarrollo evolutiva que todas las personas, en todas las culturas, debieran transitar por igual. Un proceso que permite a las personas debatir diferentes concepciones acerca de lo que es bueno para, finalmente, alcanzar un sexto estadio de madurez cuyo contenido es el de los derechos y la justicia. En palabras de Kohlberg, implica «una continua diferenciación entre la universalidad moral y los hábitos y creencias más subjetivos o culturalmente específicos» (1992:282).

Ahora bien, cuando se introduce la cuestión de la diferencia sexual en la ecuación, ya resulta imposible reducir la respuesta a una única verdad. En esa línea, Gilligan es contundente al afirmar que *las mujeres piensan diferente*. La diferencia es evidente también en lo que hace a las conclusiones normativas de sus procesos de razonamiento, pues al desarrollar el modelo ético del cuidado —contrapuesto con el de la justicia— Gilligan vislumbra la posibilidad de un mundo plural.<sup>2</sup> Uno en el que las diferencias entre concepciones morales, más que culturales, tienen que ver con el género.

Para describirlo mejor, Gilligan retoma datos de las experiencias relacionadas en las entrevistas y evalúa el modo en que estas van delineando las diferencias mencionadas. Por ejemplo, reconoce que el punto de partida de hombres y mujeres en el desarrollo no es el mismo, así como tampoco lo es el progreso hacia la elaboración de conceptos y juicios formales. Como se dijo antes, las mujeres transitan una infancia/adolescencia en la que tensionan el sacrificio por los demás y los reclamos de equidad y derechos, hacia un entendimiento de las relaciones humanas que se enlaza con una ética del cuidado. En el caso de los hombres, los absolutos de la verdad y la justicia son cuestionados con la aparición de conflictos de intereses y visiones opuestas (Gilligan, 1993:166). Las expectativas y desafíos del entorno exigen desarrollar vías de resolución que van complejizándose con la madurez mental; de allí que el entendimiento que hombres y mujeres alcanzan es verdaderamente disímil.

Un dato destacable en el análisis es la resistencia de las mujeres a juzgar opiniones y prácticas ajenas desde el propio punto de vista (Gilligan, 1993:102). Aunque, según Gilligan, esto no constituye un signo de relativismo, ni mucho menos de déficit en los procesos mentales femeninos. Tampoco nos conduce hacia un mundo en el que la moral es un amasijo de creencias y emociones dotadas de cierta estructura. Por el contrario, es una práctica que llama la atención sobre la ocurrencia de pensamiento moral complejo: el estadio posconvencional, caracterizado por un modo de entendimiento moral más completo e integrador. Uno que incluye la comprensión de «la determinación social y psicológica del comportamiento humano, así como una reafirmación de la preocupación moral en el reconocimiento del dolor y sufrimiento como realidad humana» (103); que resuelve las problemáticas en el reconocimiento tanto de la justicia y el cuidado; de acuerdo con una estructura evaluativa que funciona «según la manera en que el problema es encarado» (167).

---

2 Decimos «plural» porque, según Gilligan ambos modelos éticos son complementarios. Expresamente, resalta la necesidad de que tanto la integridad como el cuidado sean incluidas en una moralidad que permita a las mujeres enfrentar los dilemas que les impone la vida adulta, dando lugar también al reclamo de equidad y derechos. Por su parte, la ética de los deberes masculina se ve limitada en diversos dilemas, lo cual torna necesario inducir una nueva idea de responsabilidad y cuidado (1992:166).

Así, la búsqueda de una teoría moral sólida encuentra en la propuesta de Gilligan una alternativa que elude el formalismo de las reglas universalmente aplicables que caracteriza a la posición de Kohlberg. Concluye en un estadio posconvencional que relata los conflictos y soluciones morales desde la «voz femenina», antes atribuida a un nivel de desarrollo secundario o relegado al ámbito de lo privado y las emociones.

### **¿Cómo se formulan las verdades morales?**

La irrupción del pensamiento kantiano en la filosofía instala la idea de que la moralidad es un conjunto de principios racionalmente justificados, caracterizados por cierta pretensión de universalidad (Tronto, 1987:12). En esa línea, Kohlberg acuerda que la «principalidad» es la norma que gobierna el juicio moral (1992:292 y ss.). En el sentido clásico mencionado, esto supone aceptar que la moral se formaliza en reglas generales que rigen pese a las circunstancias históricas cambiantes y las arbitrariedades del contexto. Sin embargo, Kohlberg no reconoce que los principios sean entes rígidos o carentes de excepciones; pues reconoce que pueden ser interpretados para construir soluciones ante problemas prácticos dados.

Lo antedicho se ordena con otros rasgos de la teoría ya mencionados: el supuesto de la universalidad y con la idea de que es posible ofrecer una justificación racional e imparcial de dichos principios, así como de su validez general. Asimismo, con una concepción «constructivista» de la moral que concibe a los principios como el resultado de una cimentación evolutiva, más que como axiomas que dictaminan resoluciones a priori.

Tal es así, porque Kohlberg defiende un sexto estadio moral cuyo proceso de toma de decisión es «de rol ideal (...) en el que cada persona en su imaginación cambia de lugar con cada una de las personas del dilema antes de establecer su postura como la correcta» (1992:295). Fundar una decisión en semejante posición es compatible con la existencia de principios morales, pero también con un modelo de moralidad que permite a las personas evaluar situaciones concretas y perspectivas; con la idea de que los principios son contruidos y a la vez testeados en la confrontación con los casos reales. Podemos figurarlo como un espiral en el que en un extremo encontramos juicios morales, y en el otro aquellos que resultan mejor elaborados y justificados.

Como alternativa, las posiciones contextualistas defienden que la moralidad debe ser situada concretamente, eludiendo la elaboración de reglas formales y absolutas. En línea con esto, Gilligan entiende que las respuestas existen para un caso concreto, actores específicos, y se hallan insertas en una

sociedad determinada. Al concentrarse en actos singulares, sus soluciones dependen también de las capacidades de las personas implicadas. Por ello, algunos sostienen que un modelo metaético contextualista nos debe ofrecer una explicación completa del «yo», deslizando el interrogante moral fundamental «¿cuáles son los principios mejores?» hacia la preocupación por «¿cómo mejor equipar a los individuos para actuar moralmente?» (Tronto, 1987:12-13). Por ello, para Gilligan se torna central la definición de las virtudes femeninas —el cuidado, la responsabilidad, la bondad— y la voz moral que de ellas emerge. Tal punto de vista permite interpretar detallada y contextualmente un problema, sin los límites que impone la vigencia de principios —a veces, considerados arbitrarios.

Este contraste entre ambas posturas no es menor, pues la cristalización de las consideraciones morales a modo de principios generales delinea los procesos de razonamiento que resolverán los asuntos morales. Liberar la moralidad de la rigidez principalista permite dirigir las decisiones y la atención hacia los aspectos variables de los problemas; a lo que exigen las relaciones y los sentimientos hacia los demás, más no al respeto por juicios que han sido premeditados para agentes ideales, que mantienen relaciones abstractas o se preocupan por situaciones hipotéticas.

Sin embargo, el reconocimiento de la diferencia en la construcción de la identidad y la moralidad que Gilligan emprende no es gratuito; pues supone renunciar a la imparcialidad y estabilidad que garantizan los principios. Conlleva problemas suscitados, entre otras cuestiones, por la existencia de relaciones especiales y la inclusión. Por ejemplo, planteando interrogantes: ¿debemos mayor cuidado a las personas con quienes guardamos una relación especial? o ¿cuán amplia debiera ser la red de cuidado? (Tronto, 1987:14-15).

Asimismo, se acusa que tampoco un formalismo riguroso revela un enorme compromiso con las personas y la sociedad. ¿Por qué agentes ideales, desligados de sus comunidades y experiencias, habrían de sentirse obligados con otras personas? ¿Por qué se preocuparían por su comunidad o por conservar las relaciones que les atraviesan? Por eso Kohlberg considera que los principios morales son flexibles o, mejor dicho, reglas que admiten excepciones cuando se los interpreta a la luz de las circunstancias y el caso concreto. Esto revela, por un lado, cierta *sensibilidad ante la vida real*, que resulta ser la principal preocupación de quienes suscriben posiciones relativistas y contextualistas. Pero también, una evidente *estabilidad de principios*, que permite superar cualquier alternativa relativista (Kohlberg, 1992:296).

Estas diferentes explicaciones contrastan más aún cuando se piensa en las instituciones sociales y políticas en las que se insertan. Son las ventajas de ambas posturas si se considera el modo en que pueden ofrecernos respuestas

para una mayor y mejor comprensión de los problemas reales, compatibles con un régimen de derechos y deberes, reforzando también la preocupación por mantener, no solo los vínculos sociales y personales que le preocupan, sino la comunidad de pertenencia.

### **¿Qué significan los términos de la moral?**

En lo que hace a la cuestión lingüística, Kohlberg entiende que los juicios morales tienen un sentido prescriptivo; es decir, dirigen o guían la acción humana (1992:285 y ss.). Esto significa que sus enunciados suponen un compromiso particular del agente con aquello que se expresa, así como un uso concreto ligado a la propia acción.

El debate se suscita con las teorías metaéticas llamadas «descriptivas» que asignan a los juicios morales un significado particular: el de describir hechos del mundo de diversa naturaleza. Entre ellas, Kohlberg fundamentalmente discute las concepciones «naturalistas», que asignan a los enunciados morales la particularidad de describir hechos naturales del mundo que pueden ser verificables empíricamente. De forma tal, que aquello que sucede o *es*, concluye en una verdad moral.

Esto último, llamada «falacia naturalista», es el tipo de argumento que Kohlberg procura discutir. Para él, aunque el entorno constituye un estímulo que promueve el desarrollo de capacidades cognitivas y morales, de ningún modo identifica a los juicios morales con los hechos llanos del mundo natural.

La respuesta de Kohlberg a una pregunta metaética psicológica fundamental —*¿cómo se forman nuestras preferencias morales?*— contribuye a comprender lo anterior. Dicho interrogante supone llamar la atención en los elementos que se concatenan para conformar el grueso de preferencias y/o principios que rigen la moralidad; y, con ello, el rol que desempeñan los hechos del mundo y el entorno, las emociones, las creencias, etcétera.

Sin dudas que la comprensión de los seis estadios morales que Kohlberg describe está atada a la consideración de las características del entorno físico en el que se desarrollan las personas (instituciones sociales, familia, etc.), así como a las relaciones entre las experiencias que los niños y niñas atraviesan con las estructuras de sus conductas. Tal es así porque dicho desarrollo es el resultado de la interacción con aquello que les circunda. Concretamente, Kohlberg entiende que las estructuras de pensamiento se configuran en función de la adaptación del orden moral a los mandatos básicos de la cultura y prohibiciones que forman parte del tejido social, todo lo cual permite avanzar de un estadio de desarrollo a otro (1992:96-97).

Sin embargo, no por eso el desarrollo y el aprendizaje son un mero reflejo del entorno; mucho menos el resultado de experiencias pasadas. Por el contrario, para Kohlberg los criterios morales de corrección o incorrección son el resultado de un ejercicio racional práctico que incluyen a todas las anteriores; pero, fundamentalmente, implican un compromiso del hablante con los criterios o reglas en cuestión, así como la aceptación de que dichos criterios son universalizables (1992:286).

En esa línea, postula una versión prescriptivista de los asuntos metaéticos que reconoce a los enunciados morales la potencialidad de provocar una reacción en el interlocutor o interlocutora (de sujeción u obediencia). De acuerdo con esto, el juicio «no matar» es percibido como uno que prescribe que «no deberías matar». Aunque, agrega Kohlberg, los juicios morales son más que simples órdenes o prescripciones; más bien, implican un compromiso con la vigencia de principios cuya validez universal no es rigurosa en sentido sustantivo fuerte —kantiano— sino en un sentido lógico. Así, quien sostiene un enunciado moral se compromete con la vigencia de dicha regla para cualquier otra situación que posea las mismas propiedades, mas no con el presupuesto de que «todo el mundo debe permanecer fiel a las reglas universales y gobernar toda su conducta de acuerdo con ellas, o que uno no debe hacer excepciones en provecho propio» (Hare en Kohlberg, 1992:287).

Por su parte, tampoco Gilligan descarta que el entorno juegue un importante papel en la construcción de la identidad y los juicios morales. Expresamente reconoce que el conflicto, las experiencias en el mundo, etc. constituyen el motor de crecimiento y desarrollo humano. Asimismo, su énfasis en los datos biológicos humanos al definir modelos de análisis no hace otra cosa que reiterar cuán determinante es la naturaleza para la elucidación de verdades morales. Inclusive, algunos (Grimshaw, 1995:661) sostienen que, para Gilligan, ciertos antecedentes de la práctica social pueden generar prácticas morales concretas; y, con ello, una visión acerca de lo que es bueno, intereses y prioridades que deben ser respetados, o concepciones acerca de la virtud.

De todos modos, no puede ser tan simple acusarle a Gilligan una base naturalista por su mero reconocimiento del entorno en la configuración de la mente. En especial tratándose de una teoría cuya ventaja ha sido la de reaccionar contra las estructuras de las teorías tradicionales, construidas en torno a una comunidad de la cual las mujeres fueron siempre excluidas.<sup>3</sup> De hecho, Gilligan saca a la luz la injerencia que las instituciones sociales y políticas tienen en

---

<sup>3</sup> Decimos esto porque un reclamo feminista ha sido la vigencia de estereotipos de género que asigna a las mujeres no solo roles, sino también virtudes concretas, siempre vinculadas con el cuidado personal y privado, y la vulnerabilidad. Si esto es así, cualquier construcción de verdades morales en torno al contexto y aquello que meramente es, no puede alegar neutralidad o imparcialidad.

la construcción de la identidad y las motivaciones; aunque los estadios de desarrollo moral que describe no nieguen la experiencia de los grupos relegados.

Tampoco Gilligan cree que los enunciados morales procuren la sumisión u obediencia de los pares. Para ella, no son juicios que asumen universalidad lógica o normativa, ni recurren a estructuras formales rigurosas que podamos pacíficamente suscribir.

Así, la indagación por el significado de los juicios morales para Gilligan nos lleva a considerar una posición alternativa que alude a los juicios morales como expresiones de actitudes o emociones de las personas ante los hechos del mundo. De hecho, no son pocos los análisis que reducen la discusión entre Kohlberg y Gilligan a uno acerca del papel de la razón y las emociones en los procesos mentales. Desde esa mirada, mientras el constructo de los juicios morales masculinos emerge como el resultado de un análisis lógico racional acerca de principios cuya jerarquía y prioridad debe ser evaluada en función de las características de un caso; la moralidad propiamente femenina aparece vinculada con emociones —como la empatía, la bondad o el apego— que motivan las conductas y decisiones.

Sin embargo, resulta injusto acusar tal emotivismo desnudo a Gilligan. Pues supondría negarle el costado constructivista y la racionalidad que estructura todo su modelo de razonamiento moral. Al igual que en la teoría de Kohlberg, no es sino el razonamiento el que, estimulado por los dilemas del contexto, se va complejizando y abriendo paso a una forma de definir las mejores razones para justificar un determinado camino de acción. En definitiva, según Gilligan, el razonamiento moral es un ejercicio práctico; no un mero reflejo de nuestras emociones. Es el que permite avanzar en los diferentes estadios de desarrollo hacia una instancia de madurez integral.

### **¿Qué tipo de ejercicio mental supone el razonamiento moral?**

De la mano del desarrollo de las capacidades cognitivas y morales, surge la pregunta por los procesos de formulación de las teorías morales y por el razonamiento moral en sí. Esto supone indagar acerca del tipo de ejercicio práctico que dicho razonamiento implica o la forma en que guían la elección y acción individual. Es decir, «por “el modo en que razonamos acerca de lo que debemos hacer” y no tanto por aquello “que debemos hacer”» (Richardson, 2018).

Según Kohlberg, el desarrollo del razonamiento moral transita una secuencia que está íntimamente vinculada con múltiples elementos del entorno. Por un lado, con el aprendizaje y el desarrollo de las capacidades cognitivas. Por el otro, con la configuración de la identidad y las experiencias pasadas. Como se



sostuvo, es un proceso evolutivo que presupone que las personas son agentes racionales, pero, fundamentalmente, agentes que atraviesan experiencias específicas en el crecimiento.

Pero el razonamiento moral es mucho más que el resultado de dichos factores. Supone una estructura según la cual los juicios morales pueden ser verdaderos o falsos, pero no por ser reflejo de hechos, sino por el tipo de operaciones racionales que pueden derivar en ellos. Según Kohlberg, el respaldo de los enunciados morales en razones construidas en torno a la deliberación de principios explica la adecuada relación que existe entre los juicios que una persona emite con sus compromisos y conductas. No tanto así con sus opiniones y emociones —cambiantes— o con aquello que observa, arbitrario a veces.

Esto último, que se ha dado en llamar la toma del «punto de vista moral», revela el formalismo que caracteriza a los juicios morales en los que piensa Kohlberg (1992:288–289). Dicha posición reconoce la posibilidad de estar de acuerdo con los demás al definir la moralidad, sin necesariamente acordar el contenido de un conjunto de principios. Ya que implica la adopción de una posición no egoísta —llámese ideal— de compromiso con la posibilidad de que las máximas asumidas tengan validez universal. De allí que la justificación formal, lógica y racional de la moralidad resulte más relevante y plausible que su validez empírica o, inclusive, normativa.

Kohlberg contrasta este formalismo metaético con la distinción forma–contenido (1992:290 y ss.). De acuerdo con ella es posible para alguien suscribir un conjunto de principios morales sobre los que otras personas pueden estar de acuerdo sin que necesariamente consensuen su contenido. En sus palabras:

Hemos dicho que el juicio moral es prescriptivo, no es simplemente la expresión de actitudes y emociones. Hemos dicho que es cognitivo y que da razones. Sin embargo una definición más completa de un juicio moral se basa en analizar sus propiedades formales. (1992:290)

Así se aclara el papel que el razonamiento, entendido como la capacidad de ofrecer razones acerca de respuestas y enunciados, desempeña en la formulación de teorías morales. Estas resultan de construcciones que las personas elaboran en función de la ponderación de principios e intuiciones básicas.

Esto marca con claridad las posiciones de las que Kohlberg pretende tomar distancia. Por un lado, del empirismo, ya que Kohlberg no concibe a la moralidad como el resultado o las conclusiones de las experiencias. Por otro lado,

de la mera aplicación axiomática de principios rígidos.<sup>4</sup> Más bien, concibe una estructura «constructivista» de la moralidad que aúna tanto la evaluación de principios o verdades morales, con reacciones humanas fundamentales ante dichos principios y dilemas. Así, el «punto de vista moral» es la capacidad de justificar o validar principios para todo aquel que tenga «la mente clara y lógica y sabe todo lo que le concierne a él mismo, a la humanidad y al universo» (1992:291).

Sin embargo, esto es solo una parte de la historia de desarrollo moral. Gilligan muestra que introducir en el análisis datos sobre la vida y las percepciones femeninas permite informar otra descripción acerca del desarrollo y, concretamente, de las personas. No ya la de agentes que reducen sus dilemas a la lógica y la autonomía, sino que perciben las múltiples conexiones que les vinculan con sus pares, y el modo en que estas pueden verse comprometidas cuando sucede un conflicto.

La reticencia de Gilligan a elaborar una teoría universalista y formalista en un sentido duro, o la renuncia al apriorismo riguroso, pudiera conducirnos a señalar que se trata de un argumento emotivista, o uno de corte relativista, que justifica toda una estructura de principios y razones en los acuerdos circunstanciales que las personas pudieran tener. Tal es el camino que algunas críticas a su teoría han elegido.

Sin embargo, no debe perderse de vista el elemento racional que estructura el desarrollo para Gilligan. Por un lado, la referencia a las emociones que moldean el razonamiento y el crecimiento no reducen el desarrollo y la moralidad a las emociones puras. Tampoco el anclaje del desarrollo moral en la experiencia y los hechos es signo de una formulación naturalista. Más bien, lo anterior pone en evidencia el modelo de razonamiento que está implícito en el argumento.

En esa línea, Gilligan considera que las personas pueden articular sus preocupaciones e intereses ajenos en el marco de un dilema moral concreto, y ordenarlos en función de las necesidades y relaciones implicadas, sin necesidad de acudir a alguna estructura externa de principios que anticipe soluciones o anteponga límites al razonamiento. Se trata de una posición constructivista que, a diferencia de la de Kohlberg, se desprende del razonamiento en torno a principios para reflexionar bajo la lógica de las relaciones.

---

4 Aunque, como se sostuvo antes, esto no implica que Kohlberg renuncie a la vigencia de principios morales, ya que las personas requieren de ellos para la interpretación y resolución de dilemas concretos.

## **¿El desarrollo de aptitudes morales nos conduce hacia verdades morales?**

¿Existe cierto paralelismo entre el desarrollo de las capacidades morales con la formulación de verdades morales? ¿La capacidad de evaluar dilemas, ponderar valores y aplicar una respuesta a un caso real van de la mano del reconocimiento de principios morales? ¿Ciertos estadios maduros de desarrollo nos conducen linealmente hacia verdades universales?

Sin dudas que tanto Kohlberg como Gilligan asignan a la moralidad una función social vital, pues ofrece soluciones para los múltiples condicionamientos que plantean el entorno y sus desafíos. Asimismo, ambas teorías son muy claras al marcar la distinción forma-contenido que, como se mencionó, supone la posibilidad de acordar respuestas concretas y formales a los diferentes dilemas, sin que sea necesario estar de acuerdo acerca de su contenido (Kohlberg, 1992:291-292). En cuyo caso, es necesario trazar la distinción entre el desarrollo del razonamiento moral de la práctica del razonamiento y el contenido de las teorías morales o, mejor dicho, el resultado del razonamiento moral.

En este punto, tanto la teoría de Kohlberg como la de Gilligan ensayan neutralidad, pues construyen una estructura respecto del desarrollo moral y el razonamiento, que no compromete el contenido normativo de la respuesta.

Sin embargo, los principios morales que suscriben constituyen una instancia del desarrollo moral; concretamente, del último estadio. Tanto la justicia como el cuidado auguran la respuesta a los problemas morales que se presentan. En el caso de la teoría de Kohlberg, el pensamiento posconvencional implica la capacidad de defender principios en función de razones; pero también la aceptación de una posición teórica concreta —la justicia—. En el caso de la propuesta de Gilligan, las éticas del cuidado ofrecen una mirada sobre la moralidad alternativa a la clásica, pero que tampoco deja de constituir un conjunto uniforme y sólido de respuestas justificadas.

Probablemente este sea uno de los puntos en los que convergen las críticas a ambas posiciones. A Kohlberg se le ha señalado que las personas en el sexto estadio no podrían suscribir diferentes concepciones acerca de la moral (Thomas, 1995:631-632). Problema de que Gilligan no puede despegarse fácilmente ya que, aunque la justicia deja de ser el valor predominante, «el modo femenino de pensar» acerca del cuidado es el juicio moral que se impone.

Así, el desarrollo moral, una secuencia lineal que sirve para desplegar y mejorar la capacidad de pensar acerca de los problemas se convierte, en realidad, en el proceso de diálogo de la persona con teorías diversas que, finalmente, asumen una última concepción deontológica concreta.

Esto último da pie para mencionar otros problemas que suscitan los presupuestos normativos de una y otra posición. Como se señaló, la propuesta de Kohlberg sobre la justicia perpetúa un discurso androcéntrico acerca de la moralidad y el contexto. Pero Gilligan, reduce la diferencia moral a la diferencia sexual, excluyendo otros elementos del mundo que pudieran ser también relevantes. ¿Por qué no incluir también a la educación, la ocupación o la proveniencia de clase en la ecuación? Si el género y su percepción cultural inciden en el desarrollo moral, ¿cómo no es posible que otros factores sociales y culturales sean también determinantes? (Boldizar, Wilson y Deemer, 1989).

En esa línea, no debemos pasar por alto otras importantes críticas realizadas a la concepción binaria de la sexualidad (y de la moralidad) de Gilligan, que excluye a minorías y disidencias. Asimismo, se le cuestiona justificar el contexto social y político: un contexto que ha funcionado siempre bajo la mirada masculina occidental, asigna privilegios en función del género (y la raza, la religión, etc.) e identifica a la mujer con el valor de la bondad, el sacrificio, la responsabilidad y el cuidado, relegándola al mundo de lo «privado».<sup>5</sup> En definitiva, a una concepción hegemónica y androcéntrica acerca de la persona, algunos acusan que la ética del cuidado contrapone otra posición igualmente opresora (Anderson, 2020).

## Consideraciones finales

A lo largo del capítulo se explicitaron los presupuestos filosóficos de las teorías acerca del desarrollo moral que defienden Kohlberg y Gilligan. En cuanto a los fundamentos metaéticos de Kohlberg, se ha argumentado acerca de la universalidad y formalismo que adoptan sus principios, el rasgo prescriptivo de los juicios morales, y el cognitivismo y constructivismo que caracteriza al razonamiento moral. En el caso de la teoría de Gilligan, se refirió al pluralismo y contextualismo que ordena su estructura, debatiendo en torno a otras posiciones con las que se la ha emparentado, como el relativismo, el naturalismo y el emotivismo.

En dichos apartados se hizo evidente el papel que el género tiene en el desarrollo moral. Frente a una concepción predominante androcéntrica, Gilligan contrapone una que reconoce los valores y modos de pensar femeninos. Esto tiene efectos no solo en relación con las estructuras cognitivas y psicológicas que ordenan el razonamiento, sino también con los procesos mentales y los juicios prácticos resultantes.

---

5 Véase: Crenshaw (1991), MacKinnon (1989).

Sobre el final, se mencionaron algunas discusiones a las que dio lugar la contraposición entre las éticas del cuidado y las de la justicia. Por ejemplo, el modelo de moralidad que Gilligan saca a la luz ha sido fructífero en el desarrollo de éticas aplicadas particulares, en la discusión por su utilidad para los asuntos políticos y la esfera pública o, inclusive, se ha discutido si la diferencia sexual hace a mejores comportamientos éticos.<sup>6</sup>

Pero también, ha recibido numerosas críticas que no pueden soslayarse. Ya que reducir la diferencia en el desarrollo a las categorías del género o pensar a la naturaleza masculina/femenina como algo monolítico no deja de encubrir las opresiones domésticas y sociales que tradicionalmente han moldeado los valores «feminizados». En definitiva, que las mujeres tengan una aptitud natural para el cuidado de las relaciones y personas que le rodean, así como cierta propensión a evitar el daño y las resoluciones egoístas, no es otra cosa que un estereotipo sesgado por una mirada masculina del mundo. Y, fundamentalmente, una visión binaria de la naturaleza y el desarrollo moral que silencia las voces de grupos y disidencias que no se identifican con tales categorías.<sup>7</sup>

---

6 Véase: White (1999), Nodding (1984, 2013).

7 Sobre estos temas, véase: Norlock (2019), Crenshaw (1991), MacKinnon (1989), entre otras.

## Referencias bibliográficas

- Alfano, M.; Loeb, D. y Plakias, A. (2018). Experimental Moral Philosophy. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2018/entries/experimental-moral/>
- Anderson, E.; Willett, C. y Meyers, D. (2020). Feminist Perspectives on the Self. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/spr2020/entries/feminism-self/>
- Boldizar, J.P.; Wilson, K.L. y Deemer, D.K. (1989). Gender, life experiences, and moral judgment development: A process-oriented approach. *Journal of Personality and Social Psychology*, 57(2), 229.
- Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color. *Stanford Law Review*, 43(6), 1241–1299.
- Gilligan, C. (1993). *In a different voice: Psychological theory and women's development*. Harvard University Press.
- Grimshaw, J. (1995). La idea de una ética femenina. En Singer, P., *Compendio de Ética* (pp. 655–666). Alianza.
- Kohlberg, L. (1992). *Psicología del desarrollo moral*. Desclée de Brouwer.
- MacKinnon, C. (1989). *Toward a Feminist Theory of the State*. Harvard University Press.
- Noddings, N. (1984). *Caring: A Feminine Approach to Ethics and Moral Education*. University of California Press.
- Noddings, N. (2013). *Caring: A Relational Approach to Ethics and Moral Education*. Second Edition. University of California Press.
- Norlock, K. (2019). Feminist Ethics. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2019/entries/feminism-ethics/>
- Rachels, J. (2006). *Introducción a la filosofía moral*. Fondo de Cultura Económica.
- Richardson, H.S. (2018). Moral Reasoning. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2018/entries/reasoning-moral/>
- Superson, A. (2020). Feminist Moral Psychology. En Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/feminism-moralpsych/>
- Thomas, L. (1995). La moralidad y el desarrollo psicológico. En Singer, P., *Compendio de ética* (pp. 621–635). Alianza.
- Tronto, J. (1987). Más allá de la diferencia de género. Hacia una teoría del cuidado. *Signs: Journal of Women in Culture and Society*, 12(4), 644–663.

## 22. El ideal de pureza en la configuración de la identidad moral: conciencia e intersubjetividad desde el pensamiento de Hannah Arendt

*Rebeca Canclini*

Schrecklich ist die Verführung zur Güte.  
*Bertolt Brecht, Der kaukasische Kreidekreis*

El vínculo entre la bondad moral de la persona y su participación pública es un tema recurrente de la filosofía práctica desde la época de los sofistas. Sin embargo, es en el contexto de la disolución de las comunidades políticas que se consolida entre fines del siglo XVIII y principios del XIX que la reflexión filosófica sobre esta problemática se intensifica. Este escrito hace foco en la construcción de la conciencia moral en el ámbito de sociedades fragmentadas que demandan individuos capaces de participar en un mundo plural. Se propone como objetivo principal la caracterización y delimitación conceptual de un tipo que llamaremos «moralista» a partir de algunos aportes teóricos de Hannah Arendt. En general, el tipo moralista opera con criterios de limpieza y de ausencia de mezcla con elementos que se consideran perniciosos. Este tipo subvierte lo moral mediante la combinación de dos estrategias complementarias: absolutizar un solo valor o principio e identificarse plenamente con él bajo el ideal de pureza.

A partir de la figura del espectador de Hannah Arendt, se analiza esta peculiar conformación identitaria en dos niveles: el de la interioridad y el de las relaciones intersubjetivas. En este escrito se sostiene que la conformación identitaria individual tiene una relación simétrica con las formas de interacción social vigentes. Se señala que el desdoblamiento interior que da lugar a la conciencia moral se ve desfavorecido por las dinámicas sociales contemporáneas. Por otra parte, la aparición pública del moralista con su pretensión de pureza moral acentúa la fragmentación social ya existente.

Aunque algunas características del tipo moralista han sido reconocidas de manera aislada por distintos pensadores de la tradición filosófica,<sup>1</sup> las figuras del inquisidor kantiano y de la Antígona de Hegel son especialmente relevantes como antecedentes filosóficos.

---

1 En Teofrasto, por ejemplo, se presentan distintos caracteres morales que comparten algunas características con el tipo moralista (Teofrasto, 2007).

En *La religión dentro de los límites de la mera razón* Kant presenta la figura del inquisidor que condena a muerte a una persona acusada de incredulidad (1981:182). El planteo kantiano gira en torno a la conciencia del inquisidor como juez capaz de «conocer los corazones». Kant descarta que la acción de condenar a muerte a la persona se trate de falta de conciencia por parte del inquisidor. Más bien, sostiene que para actuar con arreglo a la propia conciencia, el juez debería poder conocer con certeza la voluntad del acusado, o sea, el juicio de su conciencia moral.

La peculiar lectura que Hegel hace de Antígona, por otra parte, muestra a un sujeto ético que se entiende a sí mismo como autoconciencia y, por ello, actúa como un poder arbitrario. La Antígona hegeliana muestra las limitaciones de una subjetividad vacía cuya autonomía se encuentra desvinculada de todo contenido. Antígona es un sujeto ético para quien el deber no es un objeto de conciencia sino su propia esencia (Hegel, 2010:517). El rasgo fundamental de este carácter es que asume que lo que vale para ella, vale incondicionalmente. Cabe mencionar que parte del drama revolucionario moderno consiste, según Hegel, en asumir una posición unilateral que se absolutiza y margina al agente de la totalidad. Así, la Antígona hegeliana es una figura mediante la que se tematiza la irrupción de una configuración identitaria que absolutiza su propio principio en el espacio público.

En las líneas que siguen se analiza la configuración de la conciencia moral distinguiendo dos niveles de análisis: el nivel moral que tematiza la dimensión de la interioridad de la conciencia y el nivel ético que permite analizar la conformación identitaria a partir de los vínculos que se establecen a nivel intersubjetivo. También se vincula la identidad con el factor temporal involucrado en su constitución. Posteriormente, se caracterizan las sociedades contemporáneas focalizando en los escritos arendtianos sobre las sociedades pretotalitarias. Desde este punto de vista, es importante mencionar que los estudios sobre el pensamiento arendtiano tienden a focalizar el análisis en la figura de Eichmann como representativo de una configuración identitaria propia del totalitarismo nazi. Este escrito tematiza las identidades morales propiciadas por organizaciones democráticas en contextos de inestabilidad económica y fragmentación social. Por último, se distingue la pureza moralista de la inocencia, la bondad moral y la hipocresía, y se caracteriza el tipo a partir de la reducción de lo moral a un solo criterio, del rechazo a lo diferente, de su sesgo confirmatorio y alto proselitismo y, de su marcada tendencia a la victimización y la violencia.



## La configuración de la identidad moral: niveles de análisis

La conciencia moral se conforma en dos niveles diferentes: por un lado, está el plano moral de la interioridad y, por otro lado, el plano ético de las relaciones intersubjetivas. El uso que hacemos de las nociones de «moral» y «ética» no refiere a la distinción entre objeto de estudio y lenguaje (Maliandi, 2004:54) sino a la delimitación de dos momentos de la constitución de la identidad moral. La distinción entre moral y ética que comprende a la moral como objeto de estudio y a la ética como discurso disciplinar que da cuenta de esta parcela de realidad toma lo moral como algo dado. En este escrito, sin embargo, se hace foco en el vínculo entre lo íntimo y lo intersubjetivo en la conformación de la identidad personal. Por un lado, el yo moral se da en el vínculo consigo mismo en la interioridad del pensamiento. Por otro lado, el yo ético se desarrolla mediado por las interacciones personales en el mundo compartido. La distinción entre el nivel moral y el nivel ético en la conformación de la conciencia toma como hilo conductor la figura del espectador que Hannah Arendt desarrolla en sus últimos escritos.

### Nivel moral

Con claras resonancias kantianas, Arendt comprende la conciencia moral en términos de la diferencia que aloja a un otro capaz de juzgar las acciones en la propia interioridad.<sup>2</sup> En este sentido, la conducta moral lejos de ser algo dado y constitutivo de nuestro carácter, depende, sobre todo, del trato del ser humano consigo mismo. Cuando la conciencia se desdobra surge un espectador interno que, como el *daimon* socrático, permite el distanciamiento para juzgar las propias acciones (Lederman, 2016:731).

El pensar actualiza la diferencia que se da en la conciencia mediante un movimiento dialéctico, se desarrolla bajo la forma de un diálogo silencioso con uno mismo (Arendt, 2002:210, 2007:180). De ahí que la conciencia moral se presente como una voz interior que no se encarga simplemente de informar sobre costumbres vigentes ni mandamientos externos (Arendt, 2007: 90). El pensamiento es siempre peligroso, pero no por su contenido sino por su propia actividad que tiene un carácter destructivo en tanto socava los criterios y valores establecidos (175). Así, el juicio del pensamiento es negativo

---

2 Kant presenta la noción de conciencia moral como principio subjetivo de la responsabilidad por los propios actos y la distinción entre la facultad de juzgar y la facultad de juzgar que se juzga a sí misma en *La religión dentro de los límites de la mera razón* (Kant, 1986) y *La metafísica de las costumbres* (Kant, 2002).

en el sentido de que indica lo que no debe hacerse para evitar ser desaprobados por nosotros mismos (Sanabria, 2018:24). Se trata de una voz que surge cuando el yo entabla una relación consigo mismo en la soledad (*solitude*)<sup>3</sup> de la propia conciencia.

La conciencia moral, entonces, puede ser analizada considerando la singularidad del individuo quien encuentra en la soledad de su propia interioridad la diferencia y la alteridad como condiciones de la existencia del yo (Arendt, 2007:113). De esta manera, la conciencia moral se orienta a mantener la armonía del yo consigo mismo al posibilitar la coherencia entre lo que pensamos y lo que hacemos. La acción moral solo es posible si se hace uso de la facultad del pensamiento y se desarrolla la capacidad autorreflexiva del ser humano que evalúa la identidad entre el yo de la acción y el yo espectador. El espectador interior juzga las acciones contraponiéndolas con las opiniones y el discurso propios.

La importancia del ejercicio de la autorreflexión consiste, según Arendt, en su capacidad para prevenir la banalidad del mal (2002:202). Sin embargo, como veremos más adelante, la diferencia que propicia el dos-en-uno, el diálogo consigo mismo, no se desarrolla de manera necesaria por lo que «la conducta moral no va de suyo» (Arendt, 2007:85).

## Nivel ético

El nivel ético da cuenta de la dimensión intersubjetiva que interviene en la configuración de la conciencia moral. La figura del espectador que Arendt (2003) presenta como correlativa a la facultad del juicio permite vincular el yo con el mundo y con quienes lo cohabitan. El espectador tiene la capacidad de comprometerse con el mundo y desarrollar la mentalidad ampliada junto con otros (Lederman, 2016:731).

El juicio es definido como la facultad que opera con particulares sin someterlos bajo reglas generales. Se trata de una facultad crítica que permite tomar en cuenta las representaciones de los otros para compararlas con la propia perspectiva. Arendt es heredera de las tensiones internas del pensamiento kantiano.<sup>4</sup> Por un lado, sigue a Kant al considerar que lo práctico se vincula

---

3 En *La condición humana*, Arendt distingue entre *loneliness* (soledad), el estar aislado y solo, y *solitude*, la soledad que es condición para el pensar, para el despliegue del dos-en-uno, o sea, para el encuentro con el otro interior (2009:81).

4 La oposición entre lo condicionado y lo incondicionado que se encuentra en *La crítica de la razón pura*, se vincula, en la esfera práctica, con la distinción entre imperativos hipotéticos e imperativos categóricos. Así, la acción se entiende o como instrumental o como caso particular de una ley universal.

con la libertad y, por ello, la acción no puede ser concebida desde un punto de vista técnico. Por otro lado, evita considerar la acción moral como la conclusión de una deducción que parte de una ley general (Amiel, 2000:86). De esta manera, el tratamiento arendtiano del juicio da cuenta de una facultad específica que permite la conexión intersubjetiva en el mundo como coespectadores.

Arendt sostiene que:

Es más, aunque yo tengo en cuenta a otros cuando juzgo, eso no significa que amolde mi juicio al suyo. Yo sigo hablando con mi propia voz y no cuento votos para llegar a lo que creo que es correcto. Pero mi juicio tampoco es ya subjetivo, en el sentido de que yo llevo a mis conclusiones tomándome en cuenta sólo a mí mismo. (2007:146)

El juicio permite el ajuste con el criterio de realidad de la comunidad, construye un sentido común compartido. Así, el sentido de realidad depende de la diversidad de perspectivas que se dan en el mundo y, a la vez, lo conforman. Esto se realiza mediante diversas narraciones que examinan, en su singularidad, los acontecimientos y los juzgan sin criterios preestablecidos (Canciani, 2022). La facultad del juicio impide que las representaciones subjetivas puedan ser tomadas como objetivas y, a la vez, mantiene la integridad del yo. Dicho de otra manera, el dos-en-uno del pensamiento es preservado y potenciado cuando el espectador se manifiesta en el mundo de las apariencias ante los otros.

Existen, por lo tanto, dos niveles de conformación de la identidad moral: por un lado, la interioridad del pensamiento autorreflexivo mediante el diálogo que entabla el yo consigo mismo y, por otro lado, el vínculo con los otros que hace posible la mentalidad ampliada y la construcción de un sentido común (Sanabria, 2018:24). En ambos niveles, el espectador tiene un lugar central sea como «voz de la conciencia» en la interioridad, sea como testigo comprometido capaz de considerar los juicios de los otros en el mundo común y de expresar juicios provisionales que dan cuenta de su propia perspectiva.

La alteridad es esencial en cualquiera de los dos niveles. En el pensamiento, la soledad posibilita el desdoblamiento de la conciencia y el vínculo discursivo con uno mismo. En el nivel ético, la diversidad de los puntos de vista no es anulada por la mentalidad ampliada. Tanto el juicio como el pensamiento tienen relevancia pública, ya que no son posibles en el aislamiento sino en la comunicación (noción que Arendt toma de Jaspers).<sup>5</sup> La pro-

---

<sup>5</sup> Sobre la noción de comunicación en el pensamiento arendtiano: Yeatman (2011), Straehle Porras (2015).

pia dignidad humana depende de la capacidad para establecer estos vínculos tanto en la propia interioridad, como en el mundo de apariencias compartido (Quintana, 2012:440).

Cabe aclarar que la distinción entre el nivel moral y el nivel ético no debe ser comprendida en términos abstractos. Desde el punto de vista de la experiencia, los seres humanos nacen en un mundo éticamente configurado que enmarca la constitución de las conciencias morales. Dicho de otra manera, la identidad personal no se da en el vacío.

La identidad personal se configura tanto en el espacio como en el tiempo. La identidad se conforma en la experiencia compartida en el mundo común (Tsao, 2002:106) y se revela gradualmente en la continuidad del vivir (Quintana, 2010). Por ello, toda vida individual puede ser narrada (Arendt, 2009:208). Los relatos sintetizan de manera más o menos armónica la irrupción del acontecimiento en la trama de relaciones humanas y hacen posible la comprensión de la acción a la luz de la unidad narrativa. Así como la identidad narrativa oscila entre la irrupción del acontecimiento y el mantenimiento de sí, de la misma manera la identidad moral se conforma en la tensión entre la *ipseidad* y la mismidad (Ricoeur, 1996:182). Se trata de postular una noción dinámica de identidad en la que la mismidad permite el reconocimiento de la persona como la misma en cada momento de su historia y la *ipseidad* integra los cambios en la unidad del relato que se reconfigura en cada acción. De esta manera, el acontecimiento introduce discordancias en la trama que ponen en peligro la identidad como *idem* pero, a la vez, permite que la narración discorra otorgando sentido a cada cambio.

Resumiendo, la conciencia moral se construye mediada tanto por la alteridad del desdoblamiento en la interioridad y del encuentro con otros en el espacio público, como por la tensión entre permanencia y mantenimiento de sí en el orden temporal. En ambos casos, se trata de vínculos dinámicos en los que la alteridad y la diferencia desestabilizan lo dado pero permiten la construcción de un sentido común imprescindible para propiciar el diálogo consigo mismo y con los otros.

## La identidad moral en las sociedades contemporáneas

En su obra de 1820, *Principios de la filosofía del derecho*, Hegel explica que la sociedad civil es un momento de la eticidad a pesar de ser el reino de la finalidad particular.<sup>6</sup> Esta peculiar observación da cuenta, en cierta medida, de dos fenómenos contemporáneos correlativos: el alto grado de interdependencia de las sociedades contemporáneas y la proliferación de concepciones individualistas. Por un lado, en lo que refiere a la organización financiera y productiva, la interconexión llega a niveles inéditos. Por otro lado, si se considera la dimensión social, resalta su carácter fragmentario y la atomización de los individuos. Muchos desarrollos teóricos y prácticos han fracasado al intentar fundar el orden colectivo en supuestos individualistas. Esto vale tanto para propuestas de izquierda como de derecha, lo mismo para conservadores que para progresistas.

El corrimiento de los ámbitos de incumbencia del poder público que se inicia con la Edad Moderna y se acentúa a partir de la Revolución Francesa, transformó, en paralelo, el espacio privado. Cabe aclarar que lo público y lo privado contemporáneos no son equiparables a las nociones premodernas de lo público y lo doméstico. En el mundo antiguo y medieval, la esfera pública se distinguía del espacio doméstico en el que se daba tanto la reproducción de la vida como la producción. A partir de la Edad Moderna, lo público se escinde en Estado y sociedad civil (Rabotnikof, 2002:135). La esfera de lo jurídico-político es propia del Estado, mientras que la sociedad civil puede referir tanto al sistema económico-financiero como al mundo de la vida social y cultural (De Zan, 2006). El primero, tiene relevancia pública aunque es el ámbito del interés privado. En el segundo, se dan las asociaciones de adhesión voluntaria. Así, en la sociedad contemporánea, los individuos deben buscar en el espacio público la manera de satisfacer sus necesidades mediante el trabajo y, eventualmente, pueden participar de distintos tipos de agrupaciones motivados por sus intereses particulares.

Las formas contemporáneas de trabajo conducen a la escisión social y a la superfluidad de extensos sectores de la población. La alta especialización de los saberes ha desintegrado las bases para una cultura del trabajo al punto de hacer irreversible la desintegración social vinculada con la división del trabajo y la globalización de la producción y el comercio. La integración social depende de garantizar a los asalariados la seguridad en el empleo, sin embargo, esto no es posible en las sociedades contemporáneas (Gorz, 1991:91). Las

---

6 En el párrafo §182 de *Principios de la filosofía del derecho*, Hegel afirma que en la sociedad civil cada uno es fin para sí mismo, sin embargo esa finalidad solo puede ser satisfecha mediante la referencia a los otros (2004:183).

formas de trabajo contemporáneo distinguen entre una élite «privilegiada» de trabajadores especializados funcionalmente integrados y una amplia mayoría de trabajadores que no requieren cualificaciones especiales. Las empresas dependen del primer grupo al que le brindan la seguridad en el empleo a cambio de la flexibilidad funcional y amortizan las fluctuaciones coyunturales con el resto de los trabajadores. Los trabajadores no especializados pueden depender de la empresa (en tareas de oficina, vigilancia, mantenimiento, etc.) o bien, ser interinos, temporales o trabajadores a tiempo parcial. Además, las empresas tienen mano de obra fluctuante mediante subcontrataciones (Gorz, 1991:94). En muchos lugares, hay amplios sectores de la población vinculados a la economía popular que, aunque suponen redes y relaciones muy extendidas, no están en relación de dependencia (Maldovan Bonelli, 2019). Así, la mayor parte de la población está sujeta a una inseguridad estructural en tanto están integrados funcionalmente al sistema, o bien viven una existencia en la que ni siquiera tienen la certeza de ser explotados.

La alta rotación en los espacios laborales, los diversos desplazamientos poblacionales (sean o no legales), las dificultades para cubrir las necesidades básicas (incluso trabajando a tiempo completo), la deficiente sindicalización en las grandes empresas, etc. hacen que las relaciones intersubjetivas que se establecen en este ámbito no sean lo suficientemente fuertes como para construir relaciones estables capaces de dotar de sentido a la actividad. De esta manera, la escisión social conduce a la fragmentación de la vida profesional y la vida privada. Las virtudes privadas deben ir de la mano con la eficacia profesional del funcionario que, por ejemplo, es capaz de pasar del servicio a la República de Weimar al de un gobierno totalitario sin mayores conflictos morales. En cualquier caso, el reconocimiento a la identidad personal no es esperable en un mundo en el que la mayor parte de las personas son superfluas en uno u otro sentido.

A cambio de un comportamiento funcional, el trabajador espera una compensación en la esfera privada acorde con sus sacrificios. Los bienes y servicios compensatorios no satisfacen necesidades ni son de utilidad, se trata de elementos de lujo que señalan al comprador como un privilegiado. De hecho, la estratificación en clases sociales actualmente se da en razón del consumo. Por eso, el lugar que se ocupa en la jerarquía social se vincula con la posesión (y ostentación) de determinados productos.<sup>7</sup> En carácter de consumidor, el trabajador puede evadirse «del universo colectivo hacia un nicho de soberanía privada» (Gorz, 1991:67). Consecuentemente, las conformaciones

---

<sup>7</sup> Desde esta perspectiva se puede considerar, por ejemplo, el auge de los diseños digitales de firmas como Carlings o DressX.

identitarias resultantes tienden a comprender los delitos contra la propiedad como delitos contra la propia persona, incluso en sociedades en las que tradicionalmente se reconoce una función social a la propiedad. Además, en un contexto de disolución de las instituciones tradicionales, la tendencia a violentar los criterios de propiedad vigentes crece en todos los sectores sociales (Del Percio, 2009:63). Así, la fuerte carga emocional que se manifiesta ante el robo de un celular o la corrupción de algún funcionario público es proporcional al lugar social que el bien en cuestión conlleva.

Las asociaciones voluntarias de la sociedad civil reúnen a individuos que se agrupan por inclinaciones e intereses particulares comunes conformando una trama informal de relevancia pública (De Zan, 2006). Aunque estas agrupaciones pueden dotar a la vida política de una positiva dinámica democrática, también son capaces de abrir las puertas a gobiernos totalitarios. En contextos de desintegración comunal y fragmentación social, pueden cobijar a individuos desesperados por encajar a cualquier precio en cualquier parte (Candlini, 2016:82).

Con esta base social, los proyectos democráticos proponen un modelo en que individuos y agrupaciones están, al menos formalmente, en pie de igualdad en el espacio público. Sin embargo, la vida en común exige, a veces, tomar decisiones que afectan a algunas de las ideas de bien vigentes. Difícilmente el mero diálogo pueda solucionar todos los conflictos existentes, ya que se carece de una base consensual que, en las sociedades tradicionales, es fijada por la autoridad capaz de establecer los bienes comunes y jerarquizarlos.<sup>8</sup> Como la modernidad ha separado la sociedad civil del Estado, la base ética capaz de soportar un proyecto de bien común o felicidad pública está altamente debilitada.

El mundo común posible en el contexto contemporáneo, entonces, no puede proporcionar la base para el ejercicio de la prudencia porque no constituye una comunidad ética. La antigua *phronesis* aristotélica queda reducida a la habilidad (*deinótes*) para encontrar en cada situación los medios más eficaces para lograr un fin determinado, según la distinción del libro VI de la *Ética a Nicómaco*.<sup>9</sup> El carácter fragmentario de las sociedades contemporáneas, al contrario, da cuenta de valores inconmensurables e irreductibles entre sí. Así, se carece tanto de principios morales comunes como de una

---

8 En las sociedades contemporáneas, a diferencia de las tradicionales, no existe un fin o ideal de bien común. En este sentido, cabe recordar que Aristóteles afirma que no se delibera sobre los fines sino sobre los medios (Aristóteles, *EN* III, 1112b, 11-12).

9 Sobre la prudencia en Kant (*Klugheit*) y en Aristóteles ver Aubenque (1999).

autoridad capaz de establecerlos.<sup>10</sup> Esto daña la construcción del sentido de realidad que depende de la pluralidad de perspectivas en el espacio público. De esta manera, se desarrolla una concepción de la intimidad que comprende el orden moral a partir de una identidad ajena al ordenamiento social (Matas Pons, 2015:641).

Los contextos signados por la desestabilización de los puntos de referencia comunitarios potencian los procesos de desconfiguración de la identidad personal. Como ya dijimos, el ser único de cada individuo depende de la mirada de los otros, por eso, toda subjetividad es intersubjetividad (Varela, 2017). Así, el sentimiento de desposesión e impotencia tan extendido en el mundo contemporáneo propicia vínculos sociales débiles que se reflejan en identidades morales inseguras. En esta situación, la necesidad de pertenencia, de ser parte de algo mayor que uno mismo, ejerce una poderosa atracción.

### **La pureza del moralista**

La construcción identitaria moralista<sup>11</sup> debe distinguirse de la inocencia y de la bondad moral tanto como de la hipocresía: a) la inocencia es un estado anterior a la propia distinción entre lo bueno y lo malo, por ello, el moralista no es inocente; b) tampoco se trata de la bondad moral, ya que esta suele estar referida a acciones que siempre están situadas. O sea, la bondad moral parece referir a la sensibilidad que permite detectar la acción buena en situaciones más o menos complejas. Así, difícilmente sea composable con la certeza del moralista sino, más bien, con la reflexión que considera el contexto en el que se introduce la acción; c) aunque, a veces, sea difícil distinguirlos, el moralista no es un hipócrita. Por definición, el hipócrita reconoce el valor que pretende aparentar. O sea, el hipócrita instrumentaliza su aparición ante los otros pero no se autopercebe como la encarnación de la pureza moral. De hecho, como el Yago de Shakespeare, tiene plena conciencia de su falta moral.

Si se acepta que el despliegue de la vida moral permite el desarrollo de lo humano tanto a nivel individual como colectivo, el moralista constituye una subversión de la moral. La piedra angular de esta subversión está en la absolutización de un valor o principio moral por sobre cualquier otro. Las accio-

---

10 En su *Leviatán*, Hobbes sostiene que el soberano debe fijar el lenguaje común y una hermenéutica que permita que cada quien juzgue adecuadamente lo que es legítimo con el fin de evitar la dispersión de las opiniones (2009:281).

11 A pesar de la etimología de las palabras, preferimos hablar de moralismo antes que de puritanismo. El puritanismo suele estar referido al ámbito religioso, mientras que la noción de moralismo, incluye al puritanismo pero es más amplio que él.



nes del moralista se orientan a encarnar sin fisuras ese principio moral y, de hecho, es capaz de lograrlo en un grado bastante importante. El moralista no tiene una doble moral, sus acciones son coherentes con el principio o valor que postula. Sin embargo, el precio de esta coherencia es muy alto; se paga con la eliminación de la reflexión sobre el contexto en el que se actúa y de la autorreflexión capaz de cuestionar las propias certezas morales.

La violencia estructural de las sociedades contemporáneas no solamente quebranta la regla de oro kantiana, o sea, no refiere únicamente a la instrumentalización de los seres humanos con el fin de que desempeñen una función social subordinada; tiene que ver con la desobjetivización del individuo en un proceso en el que se torna irrelevante y superfluo (Bauman, 2005). Este ser y saberse desechable se da en contextos cotidianos en sociedades democráticas que tienden a la atomización social. Pero esta soledad del desarraigo no es la soledad del que piensa y dialoga consigo mismo (Kristeva, 2003:160). Se trata de un sentimiento sintomático de formas de vida en las que los seres humanos no tienen un rol estable y, por ello, está muy extendida en el mundo actual. De esta manera, predomina o una actitud conformista propia de estados de indefensión y de impotencia o bien, una actitud que busca desesperadamente la reivindicación.

La actitud conformista es el correlato del sentimiento de impotencia de ser superfluo. Esta actitud muchas veces toma la forma de un «retiro del mundo». La búsqueda de refugio en la propia intimidad es tentadora en situaciones en las que no se encuentran vías posibles de acción. Sin embargo, también puede estar asociada a individuos que se creen demasiado buenos y nobles para oponerse a un mundo tan maligno. Arendt sostiene que estas actitudes permiten evitar el mal solo en contextos totalitarios en los que, por definición, no hay vías alternativas de acción (2007:49 y ss.). En cualquier otro caso, la retirada del mundo es un acto de irresponsabilidad culpable.

La actitud que busca la reivindicación<sup>12</sup> impulsa a agruparse en el espacio público bajo distintas consignas. De hecho, las construcciones identitarias débiles otorgan una alta significación a la pertenencia a un grupo (Sánchez Sánchez, 2003a:116). Una de las características principales de estas configuraciones identitarias consiste en la absolutización de un valor por sobre cualquier otro, reduciendo por ende lo moral a un solo criterio. Además, estas agrupaciones suelen rechazar lo diferente, manifestar cierto sesgo confirmatorio y un marcado proselitismo y, tender a la victimización y la violencia.

---

12 Hablamos de búsqueda de reivindicación y no de reconocimiento porque el punto de partida de la acción moralista está en la indignación de sentir vulnerado el principio o valor moral en cuestión.

Hablando del período anterior al totalitarismo nazi, Arendt hace la siguiente observación:

Es un nosotros propio de personas sin nombre ni rostro que encuentran en este ser anónimo e indiferenciado del grupo su orgullo. La pureza fantasmagórica del grupo se intensifica cuando el resto de los lazos sociales se han perdido, la familia, la ocupación laboral regular y, con ellas, el lugar en una comunidad. Esta disposición de ánimo no muestra su peor cara hasta que el contenido se hace irrelevante ya que, en ese caso, ya desaparecieron los fines. (1998:187)

Arendt entiende que estas agrupaciones presentan el peligro, bajo condiciones muy específicas, de devenir movimientos totalitarios. Esto se debe a que acostumbran a los individuos a una obediencia ciega más que al contenido de la norma que obedecen. En el caso de Eichmann, por ejemplo, la sensación de pureza e inocencia dependió, en cierta medida, del hecho de estar obedeciendo órdenes (Arendt, 2001:92). Es importante destacar que estas asociaciones voluntarias son sensibles a tendencias totalitarias (Villa, 2006:29; Yeatman, 2011:140) aunque su desarrollo no es necesario (Arendt, 1998:152). Los individuos que participan en estas asociaciones quieren lograr ser reivindicados, o sea, independientemente de la evaluación sobre la eficiencia de sus métodos de acción, comprenden sus acciones como medios para lograr un fin. La dimensión instrumental se pierde en favor de la voluntad sacrificial del fanático en los movimientos totalitarios. Esta distinción puede ser difusa cuando se observan las situaciones particulares; sin embargo, es fundamental para cualquier proyecto político contemporáneo.

Las identidades debilitadas por las formas de socialización contemporáneas encuentran en estas agrupaciones la posibilidad de ser parte de un nosotros homogéneo que favorece la escisión según la cual todo lo bueno y justo está en la propia vereda y, en la vereda opuesta, se encuentra lo perverso, lo sucio y lo malo (Sánchez Sánchez, 2003b:161). Esta simplificación permite eliminar la confusión que genera la diversidad de criterios y puntos de vista y, a la vez, favorece la autointegración mediante un código moral binario. Así, la posibilidad de deducir la acción de la norma moral evita las incertidumbres que generan el conflicto moral propio de las situaciones concretas en las que hay que jerarquizar principios (Lariguet, 2011). Dicho de otra manera, la facultad del juicio que opera con el particular no se desarrolla. Además, la homogeneidad que se postula a nivel grupal retroalimenta la conformación identitaria individual que no da muestras del desdoblamiento que permite la autorreflexión.

La conformación identitaria moralista suele estar tan simplificada que llega a ser superficial. El ideal de pureza tiende a reducir el ámbito de lo moral al

absolutizar un valor por sobre todos los demás, abstrayéndolo de todo vínculo interpersonal. Así, la configuración identitaria se establece negativamente por oposición a todo lo que no respeta ese único principio moral absolutizado y el yo se construye entonces bajo el ideal de la pulcritud. La unilateralidad muchas veces impulsa el paso del conformismo al afán proselitista y se puede radicalizar cuando el acceso a la información es limitado y hay baja tolerancia a la alteridad. De esta manera, no es posible la mentalidad ampliada propia del juicio y se pierde el criterio de realidad que, como ya vimos, depende de la alteridad.

El discurso moralista trata a los hechos de manera poco rigurosa, muchas veces deformándolos por ocultamiento o exageración. El sesgo confirmatorio tiende a presentar lo nuevo como repetición y así se refuerza la unilateralidad. Cabe aclarar que estas actitudes están más vinculadas a una sobreestimación de la propia capacidad que a una deficiente educación formal o a una adscripción ideológica particular. Paradójicamente, en un contexto en el que se sabe superflua e impotente, la identidad moralista parece encontrar en el valor absolutizado la única base estable en su constitución. Los sentimientos de derrota y desposesión dan lugar a una visión del mundo que combina elementos victimistas con cierta tendencia a la violencia (Amiel, 2000:40). Las personalidades moralistas suelen estar más interesadas en los otros que en sí mismas. A falta de una conciencia moral con la profundidad que otorga la diferenciación en la propia interioridad, se vuelcan a mejorar a quienes ponen en riesgo su armonía reduccionista o a anularlos. De esta manera, el moralista tiene un punto de apoyo para emprender una «campana de salvación» junto con una ferviente actividad inquisidora. Se convierte en un juez insensible que descarga su ira en juicios preferentemente públicos sin considerar ninguna contraargumentación posible (Lariguet, 2013:112). La unidimensionalidad de su interioridad evita que surja la voz de la conciencia que sería capaz de fijar límites a sus acciones. De esta manera, las identidades moralistas surgen junto con formas vinculares incapaces de generar la diferenciación necesaria para la conformación de la conciencia moral.

El vínculo entre la construcción identitaria individual y grupal puede ejemplificarse con los llamados «bailes de pureza». Se trata de encuentros a los que asisten niñas y adolescentes con sus padres: ellas se comprometen a mantenerse vírgenes hasta el matrimonio y ellos a proteger la virginidad de sus hijas. Estos encuentros son promovidos por grupos que concentran sus preocupaciones morales en la administración de las prácticas sexuales de sus integrantes (y del resto de las personas). El énfasis en las vidas sexuales propias y ajenas se deduce de una visión conservadora y naturalizada de los roles sociales de varones y mujeres. Así, se propone un ideal de familia que tiene por

correlato prácticas homofóbicas, un discurso antiabortista y restringe el rol social de las mujeres a la familia en calidad de madres y esposas. Es notable que la reducción que operan en el ámbito moral elimina las ambigüedades e incertidumbres a la hora de actuar, otorga a sus integrantes un rol claramente definido y estable, y a su vez respeta la coherencia entre el principio y las prácticas. Estos bailes dan cuenta de la necesidad de dar publicidad a la pureza y, así, configurar al otro externo propiciando la homogeneidad y unilateralidad interna mediante un código binario. Por último, aunque las prácticas violentas se presentan en grados muy variables, la victimización es bastante constante y suele adquirir la forma de la defensa de sus hijos e hijas.<sup>13</sup> Se trata de identidades que se erigen como «legisladora, intérprete de la ley, fiscal, juez y verdugo» (Sánchez Sánchez, 2003a). Cabe aclarar que la participación en estos grupos no se vincula con una absoluta imposibilidad de desarrollar la conciencia moral ni con la conformación de identidades moralistas en cada uno de sus integrantes de manera necesaria. Sin embargo, propician estas construcciones identitarias. Además, estas mismas prácticas no solo se observan en agrupaciones religiosas, o sea, el contenido del principio o valor absolutizado no es determinante.

La proliferación de discursos moralistas en el espacio público potencia el aumento de identidades homogéneas centradas en valores no negociables (Mouffe, 2012:127). La liberación de los lazos comunitarios característica de la modernidad dio lugar a diversas formas de vida, pero también facilitó el surgimiento de agrupaciones que tienden a la unilateralidad. Es importante distinguir el carácter fragmentario de las sociedades contemporáneas del pluralismo. El pluralismo demanda identidades morales autónomas y autorreflexivas que se han conformado mediadas por la alteridad. En un contexto atomizado que requiere la promoción de cierto pluralismo democrático, de un sentido común capaz de alojar la incertidumbre, estos grupos tienden a profundizar la fragmentación social.

Así, la conformación social contemporánea puede ser altamente nociva para la constitución de la identidad moral (Arendt, 2007:158). La soledad contemporánea, como vimos, se vincula con la incapacidad para desplegar el dos-en-uno y da lugar a individuos incapaces de escuchar la «voz de la conciencia». Esta falta de profundidad del individuo es un espejo de las condiciones en las que se vincula con otros. Cuando el pensamiento se retira del mundo de las apariencias que es, por definición, plural, también se pierde el sentido de rea-

---

13 El movimiento social *Con mis hijos no te metas* es un claro ejemplo de la combinación entre prácticas violentas para con el colectivo LGBTIQ+ y victimización porque, mientras afirman que sus libertades están en peligro, pretenden tener derecho a sesgar la información a la que acceden sus hijos e hijas.

lidad. A diferencia de Kant, Arendt cree que es posible que el desdoblamiento característico de la conciencia moral no llegue a darse y que los individuos lleven una vida de sonámbulos (2002:214). Por ello, juicio y pensamiento pasan a ser actividades con relevancia política (Lenis Castaño, 2009:31).

## **Recapitulación**

En este escrito se sostuvo que la conformación identitaria individual tiene una relación simétrica con las formas de interacción social vigentes. Después, se señaló que el desdoblamiento que da lugar a la conciencia moral se ve desfavorecido por las dinámicas sociales contemporáneas. Así, el tipo moralista se comprendió como una configuración identitaria vinculada con situaciones de disolución de las comunidades políticas. Por último, se mostró que la aparición pública del moralista con su pretensión de pureza moral puede también acentuar la fragmentación social.

El ideal de pureza del moralista lo lleva a poner distancia con un mundo en el que hay distintas concepciones morales y, eventualmente, conflictos irresolubles. En consecuencia, el moralista mantiene su soledad al sostener un permanente escrutinio moral sobre los otros y participar de agrupaciones sociales con baja tolerancia a la disidencia. Su aversión al barro de las situaciones concretas lo hace caer en la terrible tentación de simplemente ser bueno. El moralista elige «ganarse el cielo» antes que dejar un mundo mejor tras de sí.

## Referencias bibliográficas

- Amiel, A. (1998). *Los orígenes del totalitarismo*. Taurus.
- Amiel, A. (2000). *Hannah Arendt: política y acontecimiento*. Nueva visión.
- Amiel, A. (2001). *Eichmann en Jerusalén*. Lumen.
- Amiel, A. (2002). *La vida del espíritu*. Paidós.
- Amiel, A. (2003). *Conferencias sobre la filosofía política de Kant*. Paidós.
- Amiel, A. (2007). *Responsabilidad y juicio*. Paidós.
- Amiel, A. (2009). *La condición humana*. Paidós.
- Aristóteles (2007). *Ética a Nicómaco*. Alianza.
- Aubenque, P. (1999). *La prudencia en Aristóteles*. Crítica.
- Bauman, Z. (2005). *Vidas desperdiciadas. La modernidad y sus parias*. Paidós.
- Brecht, B. (1962). *Der kaukasische Kreidekreis*. Suhrkamp Verlag.
- Canclini, R. (2016). *Un mundo para la acción. Subjetividades políticas y ley en Hannah Arendt*. EdiUns.
- Canclini, R. (2022). El concepto de espacio público en Hannah Arendt: la fuente de la metáfora teatral. Las Torres de Lucca. *Revista internacional de filosofía política*, 11(2), 323–334. <https://dx.doi.org/10.5209/itdl.77034>
- De Zan, J. (2006). Los sujetos de la política. Ciudadanía y sociedad civil. *Tópicos*, 14. <http://www.redalyc.org/articulo.oa?id=28819851005>
- Del Percio, E. (2009). *Política o destino*. Sudamericana.
- Gatta, G. (2014). Visiting or house-swapping? Arendt and Jaspers on empathy, enlarged mentality and the space between. *Philosophy and Social Criticism*, 40(10), 997–1017.
- González Vallejos, M. (2016). Kant y el inquisidor. Acerca del recurso hipotético a Dios a partir de la conciencia moral. *Anuario Filosófico*, 49(3), 565–584.
- Gorz, A. (1991). *Metamorfosis del trabajo*. Sistema.
- Hegel, G. (2004). *Principios de la filosofía del derecho*. Sudamericana.
- Hegel, G. (2010). *La fenomenología del espíritu*. Abada.
- Hobbes, Th. (2009). *Leviatán*. Fondo de Cultura Económica.
- Kant, I. (1981). *La religión dentro de los límites de la mera razón*. Alianza.
- Lariguet, G. (2009). La ética frente a los dilemas trágicos: guerreros, pacificadores y cirujanos. A parte rei, 64. <http://serbal.pntic.mec.es/~cmunoz11/lariguet64.pdf>
- Lariguet, G. (2011). Las tensiones internas del pluralismo moral. Isegoría. *Revista de Filosofía Moral y Política*, 44, 161–184.
- Lariguet, G. (2013). El aguijón de Aristófanes y la moralidad de los jueces. *Doxa, Cuadernos de Filosofía del Derecho*, 36, 107–126.
- Lederman, Sh. (2016). *The actor does not judge: Hannah Arendt's theory of judgement*. *Philosophy and Social Criticism*, 42(7), 727–741.
- Lenis Castaño, J.F. (2009). *Hannah Arendt: conciencia moral y banalidad de la condición humana*. *Co-herencia*, 6(11), 29–38.
- Maldovan Bonelli, J. (2018). *La economía popular: debate conceptual de un campo en construcción*. UMET.
- Maliandi, R. (2004). *Ética. Conceptos y problemas*. Biblos.
- Matas Pons, Á. (2015). Falseamiento y consumo de la identidad, de Rousseau a Adorno. Isegoría. *Revista de Filosofía Moral y Política*, 53, 631–646.

- Mouffe, Ch. (2012). *La paradoja democrática. El peligro del consenso en la política contemporánea*. Gedisa.
- Palomar Torralbo, A. (2010). Conformismo, banalidad y pensamiento: figuras de la alienación en las sociedades de masas según Hannah Arendt. *Astrolabio. Revista Internacional de Filosofía*, 11, 349–358.
- Quintana, L. (2010). *Identidad sin sujeto: Arendt y el mutuo reconocimiento*. *Ética & Política / Ethics & Politics*, 12(2), 430–448.
- Rabotnikof, N. (2002). El espacio público, Caracterizaciones teóricas y expectativas políticas. En Quesada, F. (Ed.), *Filosofía política I. Ideas políticas y movimientos sociales* (pp. 135–151). Trotta.
- Ricoeur, P. (1996). *Sí mismo como otro*. Siglo XXI.
- Sanabria, M. (2018). *Constitución de la conducta moral en la filosofía de Hannah Arendt: el testigo en el pensamiento y en el mundo de las apariencias*. *Perseitas*, 6(1), 23–47.
- Sánchez Sánchez, T. (2003a). ¿Cómo se fabrica un fanático? Mecanismos psíquicos en la construcción de una mentalidad fanática. *Papeles Salmantinos de Educación*, 2, 111–133.
- Sánchez Sánchez, T. (2003b). *Paradojas existencias y emocionales de las personalidades fanáticas*. *Clínica y Salud*, 14(2), 157–181.
- Straehle Porras, E. (2015). Autoridad, soledad, mundo: un diálogo entre Jaspers y Arendt. *Bajo Palabra. Revista de Filosofía*, 10, 17–29.
- Teofrasto (2007). *Caracteres*. Gredos.
- Tsao, R. (2002). Arendt against Athens: rereading «The Human Condition». *Political Theory*, 30(1), 97–123.
- Varela Manograsso, A. (2017). Las formas de la violencia y la identidad personal en Hannah Arendt. *Bajo palabra. Revista de Filosofía. II Época*, 15, 149–164.
- Yeatman, A.; Hansen, Ph. (...) y Barbour, Ch. (2011). *Action and appearance*. Continuum.

## 23. La metáfora del mal. En torno a la discusión entre Hannah Arendt y Gershom Scholem

María E. Wagon

El tema del mal es crucial para la ética y se encuentra ampliamente extendido. Dado que el tópico tiene una larga y compleja historia tanto en filosofía moral como política, a los presentes fines voy a acotar su tratamiento proponiendo para ello una reconstrucción de un hito ilustrativo y relevante: la discusión que sobre el mismo entablaron Hannah Arendt y Gershom Scholem.

El problema del mal atraviesa, explícita o implícitamente, la totalidad de la obra arendtiana, al punto que puede postularse, no sin debate, como el hilo conductor que hilvana las temáticas en las que Arendt se detiene y los conceptos que desarrolla a lo largo de su vida académica. Durante su juventud el problema del mal cobró protagonismo en sus reflexiones con el advenimiento del régimen totalitario nazi y el surgimiento de un mal sin precedentes en la historia. Estas circunstancias promovieron el abandono, por parte de Arendt, del ámbito de la especulación filosófica para incursionar en el terreno teórico político,<sup>1</sup> espacio en el que, finalmente, transitó la mayor parte de su carrera. Al final de su vida fue nuevamente el problema del mal el impulsor de la vuelta arendtiana al terreno filosófico. En la introducción de su obra inconclusa *La vida del espíritu*, Arendt explicita que lo que la motivó a dejar atrás el terreno de la teoría política para hundirse nuevamente en las complejas profundidades de la especulación filosófica fue el juicio llevado a cabo en Israel en 1961 contra Adolf Eichmann, al cual asistió en calidad de cronista.

Arendt, luego de su participación en el mencionado juicio, abandona su concepción de mal radical postulada en *Los orígenes del totalitarismo* y la reemplaza por la controvertida noción de la banalidad del mal. Dicha nueva consideración respecto del mal generó, luego de la publicación de su reporte, una controversia que aún en la actualidad se mantiene vigente.<sup>2</sup> En una carta en respuesta a Gershom Scholem, Arendt hace explícito el mencionado cambio en su concepción del mal y utiliza una metáfora explicativo-argumentativa para fundamentar su postura. Desde el ámbito crítico se ha relativizado la afirmación de la autora respecto al cambio que sufrió su concepción del mal. Por un lado, se sostiene que las nociones de mal radical y banalidad del

---

1 En la entrevista que Günter Gaus le realizara en 1964, Arendt afirma que lo que la impulsó a centrar sus reflexiones en el ámbito político fue: «the burning of the Reichstag, and the illegal arrests that followed during the same night. (...) This was an immediate shock for me, and from that moment on I felt responsible» (1994b:4-5).

2 Para una profundización al respecto cf. Ezra (2007) y Wagon (2017).



mal son complementarias en tanto responden a diferentes aspectos del mal y, en este sentido, pueden coexistir. Por otro, se afirma que la acepción del adjetivo «radical» con la que Arendt define al mal en su carta a Gershom Scholem difiere de la que utiliza en *Los orígenes del totalitarismo*. En la mencionada carta estaría tomada en sentido estrictamente etimológico, en cambio en el capítulo «Dominación total» se lo define en relación con la noción de superfluidad. Ahora bien, estas consideraciones no son aceptadas por toda la literatura pues hay quienes brindan algunas posibles razones que explicarían el abandono, por parte de Arendt, del concepto de mal radical y afirman la incompatibilidad del mencionado concepto respecto del de la banalidad del mal si se los aborda a ambos filosóficamente. Se sostiene que Arendt dejó de lado la categoría de mal radical porque era consciente del dejo teológico que dicha concepción del mal traía aparejada.

En el presente trabajo se aborda, como tema principal, el uso de la metáfora por parte de Hannah Arendt a la hora de hacer referencia al mal acaecido durante el régimen totalitario nazi. En un primer momento se realiza un comentario general sobre el problema del mal en la obra arendtiana para luego hacer hincapié en el controvertido cambio de postura realizado por la pensadora en lo que respecta a su concepción del mal. En este punto se hace mención a un intercambio epistolar entre Arendt y Scholem, en donde la autora utiliza una metáfora para fundamentar su postura así como también para explicar su punto de vista. El paso siguiente consiste en exponer brevemente las reflexiones sobre el lenguaje metafórico expuestas por Arendt en *La vida del espíritu* que servirán de herramientas para poder realizar el análisis del uso de la metáfora a la hora de comprender y poner en palabras hechos horribles sin precedentes.

## **El problema del mal en el pensamiento de Hannah Arendt. Comentarios generales**

Como se mencionara en la introducción, el abordaje del problema del mal sufre, a lo largo de la evolución del pensamiento arendtiano, una transformación innegable, cuestión que se evidencia tanto en las propias manifestaciones de Arendt (2007) como en los cambios que pueden observarse en la caracterización y definición de dicho problema. Arendt (1998b) se refiere al mal totalitario como mal radical, una noción de origen kantiano<sup>3</sup> que ella

---

3 Kant (1969) plantea tres posibles propensiones (esto es, orígenes del) al mal: la debilidad, la impureza de corazón y la malignidad (el mal radical). Arendt (1998b) sostiene que los tres orígenes del mal según Kant resultan insuficientes o inaplicables a lo acontecido en Alemania durante el régimen nazi, de ahí que resulte necesario resignificar la noción de mal radical

reformula. En tal reformulación arendtiana se puede observar que la característica principal que dicho mal presenta es su apelación a la superfluidad, es decir, su pretensión de eliminar de los seres humanos las notas distintivas de su humanidad, «ser superfluo significa no pertenecer en absoluto al mundo» (Arendt, 1998b:380).

Arendt (1998b), en el capítulo XII, apartado 3: «Dominación total», analiza el significado del mal en su vinculación con lo acontecido en los campos de concentración y exterminio nazis. Dichos campos no solo fueron creados con la finalidad de degradar y posteriormente exterminar a los seres humanos sino también con el objetivo de ser funcionales al experimento de eliminar la espontaneidad que caracteriza el comportamiento del individuo, transformando así su personalidad en una simple cosa (Arendt, 1998b). Es en esta aniquilación de la capacidad creadora y de la espontaneidad humana donde Arendt ve un mal radical. En este sentido, el mal radical surge de la mano de una nueva concepción del género humano: el ser humano como ser superfluo.

La noción de superfluidad aparece a lo largo de la totalidad de *Los orígenes del totalitarismo*. Según el análisis arendtiano, los regímenes totalitarios no pretenden lograr una dominación despótica sobre los individuos sino tornarlos superfluos debido a que esta es la única manera de alcanzar el poder total. Una de las precondiciones de dicha superfluidad es el desarraigo, es decir, convertir a los individuos en extranjeros dentro de un mundo que no los reconoce como partes integrantes. El desarraigo implica no tener en el mundo un lugar de pertenencia que sea reconocido por los otros, la superfluidad, por su parte, implica la no pertenencia al mundo.

El régimen nazi comenzó por dejar al margen de la ley a sus víctimas, aislandolos del resto del mundo por medio de la reclusión en guetos y campos de concentración. Los reclusos en los centros de detención y exterminio eran considerados muertos en vida por el resto del mundo. «The concentration-camp inmate has no price because he can always be replaced; nobody knows to whom he belongs, because he is never seen. From the point of view of normal society, he is absolutely superfluous» (Arendt, 1979:444).<sup>4</sup> En esto radica para Arendt el mal radical, un mal que nunca había acaecido en el mundo hasta la aparición del totalitarismo, un régimen de gobierno cuyo objetivo último es convertir a los seres humanos en criaturas idénticas entre sí, igualmente superfluos e incapaces de ser espontáneos (Canovan, 1992).

En la década del 60, sin embargo, la concepción arendtiana del mal realiza un giro importante que, hasta la actualidad, sigue generando discusiones en

---

4 «El internado en el campo de concentración no tiene precio, porque siempre puede ser sustituido; nadie sabe a quién pertenece, porque nunca ha sido visto. Desde el punto de vista de una sociedad normal es absolutamente superfluo» (Arendt, 1998b:356).

el ámbito de los análisis críticos. Luego del juicio al que fuera sometido Adolf Eichmann en Jerusalén en 1961, al que Arendt asistió en calidad de cronista, la pensadora acuña la noción de *banalidad del mal* para referirse al nuevo tipo de mal que surgiera de la mano de las medidas llevadas a cabo por el régimen totalitario nazi y abandona su concepción anterior respecto al mal, a saber: el mal radical. Cabe mencionar que Arendt explícitamente aclara que cuando se refiere a la banalidad del mal no pretende postular una teoría ni una doctrina sobre el mal sino reflexionar sobre una cuestión puramente factual y concreta, el fenómeno del mal, específicamente, el tipo de mal que llevó a cabo el régimen nazi en Europa (Arendt, 2003b).

Arendt (2003a) no hace ninguna alusión al mal radical ni al porqué de su cambio de postura. Para ella lo alarmante del juicio en Jerusalén fue llegar a la conclusión de que las acciones terribles por las que se juzgaba a Eichmann no fueron cometidas por un ser monstruoso, sino que, por el contrario, fueron realizadas por alguien que se encontraba muy lejos de comportarse como un demonio (Canovan, 1992). Arendt vio en la figura de Eichmann a alguien con la capacidad de realizar las peores atrocidades movido por motivos carentes de cualquier grado de malignidad, en consecuencia, llega a la conclusión de que los crímenes más perversos pueden, en muchos casos, surgir del déficit de pensamiento. Arendt (2003a) reconoce, en el «Post Scriptum» agregado en la edición de 1965 de *Eichmann en Jerusalén*, que el concepto de banalidad del mal puede resultar controvertido. La autora observa que no se puede equiparar a Eichmann con los personajes malvados de la literatura clásica en tanto no es equiparable a un Yago ni a un Macbeth, sino que es un hombre común al que la irreflexión lo llevó a cometer los peores crímenes acaecidos hasta el momento. La conclusión arendtiana es que la irreflexión puede llevar al ser humano a males inimaginables.

Ya al final de su carrera, Arendt, en *La vida del espíritu*, inicia la introducción a la primera parte, «El Pensar», con una referencia al problema del mal. La autora retoma la novedad con la que se encontrara al presenciar el juicio a Eichmann en relación con el problema del mal, pues, como se expusiera, el acusado no se mostró acorde con el estereotipo del malvado que recorre el pensamiento occidental. Los hechos que se juzgaban eran monstruosos pero el responsable de ellos era un hombre común y corriente. La ausencia de pensamiento que evidenciaban las respuestas e intervenciones de Eichmann durante el juicio fue lo que llamó la atención de Arendt. Tal ausencia no podía ser explicada por un olvido de aquellos hábitos considerados buenos ni por la estupidez, es decir, por la incapacidad para comprender.

Is evil-doing (...) possible in default of not just «base motives» (...) but of any motives whatever, of any particular prompting of interest or volition? (...) Might the problem of good and evil (...) be connected with our faculty of thought?<sup>5</sup> (Arendt, 1978:4-5)

Estas preguntas son las que motivan y guían el análisis de la facultad de pensar que realiza Arendt en su último libro. Es decir, el cuestionamiento arendtiano intenta dilucidar si es posible que el ejercicio de la capacidad humana de pensar sea una condición necesaria y/o suficiente para que las personas no cometan el mal. En otras palabras, luego de que su asistencia al juicio a Eichmann la pusiera en posesión de la idea de la banalidad del mal, Arendt se aboca a analizar la atinencia de tal expresión, sus límites y alcances, concluyendo que la irreflexión puede ser el motivo impulsor de las malas acciones, incluso de los peores horrores.

### **La metáfora del mal y el cambio de postura arendtiano**

Como se comentara anteriormente, luego de su asistencia como cronista al juicio a Eichmann en Jerusalén, Arendt abandona su concepción de mal radical y la reemplaza por la noción de la banalidad del mal. En el marco de la controversia generada por la publicación de su reporte del mencionado proceso, Gershom Scholem<sup>6</sup> le escribe una carta (junio de 1963) manifestándole su completo desacuerdo tanto respecto al planteamiento de su crónica como al tono que utiliza en algunos de sus pasajes. Por otra parte, critica el análisis arendtiano del rol ejercido por los Consejos Judíos y la relación de Arendt con el sionismo.<sup>7</sup> En lo que respecta específicamente a la noción de la banalidad del mal y al abandono de la concepción del mal radical, Scholem, en su carta del 23 de junio de 1963, escribe:

---

5 «¿Es posible hacer el mal (...) cuando faltan no ya solo los “motivos reprobables” (...) sino también cualquier otro tipo de motivos, el más mínimo destello de interés o volición? (...) ¿Puede estar relacionado el problema del bien y del mal (...) con nuestra facultad de pensar?» (Arendt, 1984:14-15).

6 Gershom Scholem fue un estudioso judío nacido en Alemania y luego radicado en Israel, se lo conoció por su erudición respecto de la mística judía. Si bien Arendt y él se conocieron gracias a un amigo en común, Walter Benjamin, ella se convirtió en una asidua lectora de sus estudios sobre el misticismo judío. Para una profundización sobre la relación entre Arendt y Scholem cf. Jorba (2019).

7 Para una ampliación de la relación de Arendt con el sionismo cf. Serrano de Haro (1998).

Tras la lectura de tu libro sigo sin convencerme de tu tesis de la «banalidad del mal» (...). Esta novedosa tesis de la banalidad del mal me suena más a un *slogan*; y desde luego no me da la impresión de que sea el resultado de un análisis tan exhaustivo como el que llevaste a cabo de manera mucho más convincente en tu libro sobre el totalitarismo, y al servicio de una tesis muy diferente, incluso contradictoria con la de este otro. (Arendt, 1998a:27)

En la carta en respuesta a Scholem (julio de 1963), Arendt ratifica el mencionado cambio en su concepción del mal:

Llevas razón: he cambiado de parecer y ya no hablo de «mal radical». (...) Mi opinión es hoy, en efecto, que el mal nunca es «radical», que es solo extremo, y que no posee ni profundidad ni dimensión demoníaca ninguna. (...) Solo el bien tiene profundidad y puede ser radical. (Arendt, 1998a:32)

La noción de mal radical ya no es retomada por Arendt, quien, salvo en la carta citada con anterioridad, no da mayores explicaciones sobre la no atinencia del adjetivo «radical» en lo que respecta al mal totalitario. Desde el ámbito crítico se ha relativizado la afirmación de la autora respecto al cambio que sufriera su concepción del mal. Por un lado, se sostiene que las nociones de mal radical y banalidad del mal son complementarias en tanto responden a diferentes aspectos del mismo mal (Bernstein, 2004; Pendas, 2007; Hilb, 2015, entre otros) y, en este sentido, pueden coexistir. Por otro, se afirma que la acepción del adjetivo «radical» con el que Arendt define al mal en su carta a Gershom Scholem difiere de la que utiliza en *Los orígenes del totalitarismo* (Bernstein, 2004). En la mencionada carta estaría tomada en sentido estrictamente etimológico mientras que en el capítulo «Dominación total» se lo define en relación con la noción de superfluidad. Ahora bien, estas consideraciones no son aceptadas por toda la masa crítica pues hay quienes (Villa, 1999; Young–Bruehl, 2004; entre otros) brindan algunas posibles razones que explicarían el abandono, por parte de Arendt, del concepto de mal radical y afirman la incompatibilidad del mencionado concepto respecto del de la banalidad del mal si se los aborda a ambos filosóficamente. Villa (1999) sostiene que Arendt dejó de lado la categoría de mal radical porque era consciente del dejo teológico que dicha concepción del mal traía aparejada. A criterio de Villa, el mal puede poseer profundidad metafísica solo en un marco teológico que postule la existencia de fuerzas transhumanas que breguen por el bien y el mal.

En la anteriormente citada carta a Scholem, Arendt utiliza una imagen metafórica para ilustrar su nueva concepción del mal y justificar su cambio en la concepción de dicho fenómeno:

Mi opinión es hoy (...) que el mal nunca es «radical», que es solo extremo, y que no posee ni profundidad ni dimensión demoníaca ninguna. Puede proliferar y arrasarse el mundo entero precisamente porque se extiende como un hongo en la superficie. «Desafía» al pensamiento (...) porque el pensamiento trata de alcanzar alguna profundidad, de ir a la raíz, y en el momento en que se ocupa del mal se ve frustrado porque ahí no hay nada. Tal es su «banalidad». (Arendt, 1998a:32)

La capacidad destructiva del mal prolifera y arrasa el mundo entero porque se extiende por la superficie como un hongo. Su banalidad radica, según Arendt, en que cuando el pensamiento lo aborda y busca en las profundidades e intenta alcanzar sus raíces, no encuentra nada. Arendt está convencida, ahora, de que solo el bien es profundo y puede ser radical. Kohn (2003) menciona que la radicalidad del mal a la que Arendt se refiere en primera instancia hacía referencia a que la raíz del mal había surgido por primera vez en el mundo. Pero lo que la pensadora evidenció durante el juicio a Eichmann fue que dicho mal tenía la capacidad de propagarse ilimitadamente por toda la tierra, como un hongo, puesto que, para hacerlo, no necesitaba nutrirse de ninguna clase de ideología.

En las lecciones que Arendt dicta entre 1965 y 1966 (2003c) sobre las cuestiones morales, retoma la idea de que el peor de los males que pueda existir no es radical, en el sentido de que no tiene raíces, por lo que tiene la capacidad de extenderse con mucha rapidez y de manera ilimitada. Cuando Arendt utiliza la noción de «raíz» o «raíces» lo hace en el sentido de profundidad y arraigo en la reflexión. El mal sin límites encuentra terreno fértil allí donde la facultad de pensar (y de recordar) está ausente y, por consiguiente, también lo están las raíces que limitan las posibilidades de obrar. Aquellos individuos que no ejercen su capacidad de pensar se dejan arrastrar por los acontecimientos, deslizándose por la superficie de los mismos, y no penetran nunca hasta la profundidad reflexiva de la que, como seres humanos, son capaces.

### **3. Analogías y metáforas. Consideraciones generales**

Arendt, en *La vida del espíritu*, afirma que todo el lenguaje filosófico y gran parte del poético son metafóricos. «Toda metáfora descubre “una percepción intuitiva de similitud en lo desemejante”» (1984:123). Por medio de una referencia a Kant, Arendt menciona que el pensar (razón especulativa en términos kantianos) solo puede manifestarse a través del lenguaje metafórico. La retirada del mundo fenoménico que funge como precondition de las actividades del espíritu, se desvanece, pues la metáfora conecta ambos mundos (el

mental y el material) sustrayendo lo abstracto del pensamiento de las intuiciones sensibles, estableciéndose, de esta manera, la realidad de los conceptos. Arendt aclara que lo anteriormente expuesto resulta sencillo en lo que respecta al pensamiento de sentido común. Sin embargo, todo cambia y se vuelve más complejo en el terreno del pensamiento especulativo. Es en este tipo de pensamiento donde la metáfora cobra vital relevancia en tanto es la que permite la traslación del pensamiento al mundo de los fenómenos, lo cual solo puede ser realizado por medio de analogías. Con una nueva referencia a Kant, Arendt afirma que las percepciones metafísicas solo se logran apelando a analogías entendidas como la total semejanza de dos relaciones entre dos realidades completamente desemejantes (1984:125).

All philosophical terms are metaphors, frozen analogies, as it were, whose true meaning discloses itself when we dissolve the term into the original context, which must have been vividly in the mind of the first philosopher to use it.<sup>8</sup> (Arendt, 1978:104)

Canclini (2016:110) retoma la cita anterior y afirma que para Arendt la analogía es la que propicia la construcción de conceptos filosóficos que se congelan y trascienden, luego, como metáforas. En el caso de que el contexto en el que fueron acuñados se explicita, es posible desvelar el significado original.

El objeto de la actividad de pensar es la experiencia y su finalidad última es comprender. Esta comprensión es definida por Arendt como pensar-sobre-una-cosa y acaece en la actualización de un pensamiento libre. El pensamiento logra salvar la distancia que lo aleja del mundo material por medio de las analogías y de las metáforas, por medio de las cuales se logra la unidad de la experiencia humana. Arendt concluye que no hay dos mundos porque la metáfora los unifica (1984:132).

## **El uso del lenguaje metafórico y el mal en la obra arendtiana**

Como se expusiera anteriormente, Arendt sostiene que el lenguaje filosófico, en su totalidad, es metafórico, pues sus conceptos adquieren realidad por medio del vínculo analógico que dicho recurso (la metáfora) establece entre el mundo fenoménico y la facultad del pensamiento. La analogía entendida como «metáfora cristalizada» es la que le otorga unidad a la experiencia huma-

---

<sup>8</sup> «Todos los términos filosóficos son metáforas, por así decir, analogías cristalizadas, cuyo verdadero significado se revela cuando disolvemos el término en el contexto originario, que tan claramente presente debió de estar en el espíritu del primer filósofo que lo utilizase» (Arendt, 1984:125).

na por medio de la relación de semejanza entre dos planos que, en su ausencia, permanecerían escindidos. Ahora bien, no obstante la afirmación arendtiana respecto al sustrato analógico/metafórico del lenguaje filosófico, la metáfora en su sentido más fuerte, es decir, en su uso poético/literario, cobra un peso mayor en lo que respecta al intento de comprender y poner en palabras el mal más extremo, el mal que advino al mundo de la mano del totalitarismo.

Heuer (2005:47) menciona que la obra de Arendt está plagada de metáforas y, por medio de una referencia a los *Diarios filosóficos*, reitera lo dicho en el párrafo anterior, a saber: que el pensamiento, a criterio arendtiano, se sirve de la experiencia, de lo visible, para formar sus conceptos y designar lo invisible. «No se puede definir la metáfora misma de otra manera que no sea por medio de metáforas» (Heuer, 2005:48). Arendt se refiere a la metáfora como el puente sobre el abismo que separa la interioridad invisible del ser humano y el mundo de apariencias. Otra imagen que utiliza Arendt en el *Diario filosófico* para referirse al uso de analogías y metáforas es la de un hilo que une el pensamiento con el mundo de la experiencia.

En lo que respecta al uso de metáforas por parte de Arendt a la hora de abordar las experiencias límite, específicamente el surgimiento de un nuevo tipo de mal en el contexto de los totalitarismos del siglo xx, Blanco Ilari (2013) menciona que, ante dichos acontecimientos, la posibilidad de la comprensión se ve puesta en jaque. Sin embargo, hay una tendencia normalizadora por parte del pensamiento que intenta abarcar y explicar los nuevos hechos mediante las categorías con las que cuenta. En esta reducción de lo desconocido al universo categorial conocido se pierde el carácter extraordinario de la experiencia, así como la capacidad humana de verse afectado por la novedad de la misma. La dificultad/imposibilidad para comprender genera conmoción, la cual se traduce en la incapacidad de nombrar los acontecimientos inéditos y extremos (327). «La *metáfora* es el recurso que tiene el lenguaje para habérselas con lo nuevo, con lo que rompe la habitualidad» (Blanco Ilari, 2013:327).

En el apartado sobre el cambio de postura arendtiano respecto a su concepción del mal se hizo mención a un intercambio epistolar entre Arendt y Scholem. En la respuesta arendtiana a las críticas de su interlocutor se ilustra, por un lado, el valor argumentativo de la metáfora en tanto Arendt la utiliza para fundamentar su postura y, por otro, el intento de poner en palabras aquello que es imposible de ser nombrado por medio de un lenguaje llano y literal. La autora se refiere al nuevo tipo de mal ante el cual se encontró durante el juicio a Eichmann, un mal que ya no puede ser catalogado de radical porque carece de la profundidad que le otorga el pensamiento. La banalidad del mal consiste, justamente, en un mal carente por completo de la profundi-



dad en la que se mueve el pensamiento. Es un mal, en cambio, que se propaga en la superficie, «como un hongo», dice Arendt, con la velocidad letal de lo desarraigado:

Pensar, escribió Arendt en su diario, siempre tira en lo que se encuentra bajo la superficie, o sea en la profundidad. «En lo profundo está su dimensión (...)». Ahí, donde «el diálogo del pensamiento... falta, ya no tiene profundidad, y es la vuelta a la trivialidad. (...) De la trivialidad resulta el desastre y no de la profundidad, ya que esta la hemos perdido (...)». (Heuer, 2005:43)

La traducción al plano de la moralidad de la metáfora del mal como un hongo sería la siguiente: el mal ingresa al mundo de la mano de un reducido grupo cuyos integrantes confluyen en una ideología común que no reconoce límites a la hora de bregar por sus objetivos, pero crece y se apodera de todos por medio de la moralidad de los ciudadanos comunes. Esto acontece debido a que, para la gran mayoría de los miembros de una sociedad en general y para Eichmann en particular, la moral se reduce a desempeñarse de acuerdo a las costumbres, actitud que promueve el cumplimiento de cualquier orden o práctica sin mayores conflictos y con buena conciencia, es decir, creyendo que su manera de desempeñarse es la correcta. «The terrible irony is that unthinking conformity joined to unexamined self-preference (self-love) suffices to produce in normal people an everyday version of the loss of human status endured in the death camps» (Kateb, 2013:355).

## **Conclusión**

Arendt, en su crónica sobre el juicio a Eichmann, habló del mal banal que sirvió de fundamento para las conductas y justificación de las mismas por parte de los diferentes agentes del nazismo. Creyó ser clara y, por medio de la noción de banalidad del mal, dar cuenta de una realidad indiscutible. Sin embargo, la publicación de su escrito generó la controversia más efervescente y extendida en el tiempo de todas. En su carta a Scholem recurre al uso de una metáfora para poder hacerse entender, para fundamentar su postura e intentar materializar en palabras aquello que mediante lo horrendo de sus consecuencias la impactó en la trivialidad de sus orígenes.

A criterio arendtiano, el mal que el mundo conoció de la mano del régimen totalitario nazi ya no tiene raíces, ni superficiales ni profundas, pues el sustrato reflexivo en las que estas podían arraigarse para crecer y fortalecerse ya no existe. La banalidad del mal remite a una nueva tipología de mal que crece y

se expande con la velocidad en la que proliferan los hongos, por la superficie. En esto radica, para Arendt, su peligrosidad, en su capacidad de replicarse y reproducirse de forma ilimitada porque no requiere, para hacerlo, de compromiso ideológico alguno. La irreflexión crea las condiciones necesarias y suficientes para que este tipo de mal se expanda y quien contribuye a aumentar su alcance y a agravar sus consecuencias, lo hace, muchas veces, con buena consciencia, creyendo que solo obedece y que no tiene, tampoco, alternativas.

## Referencias bibliográficas

- Arendt, H. (1978). *The Life of the Mind*. Harcourt, Inc.
- Arendt, H. (1979). *The Origins of Totalitarianism*. Harcourt Brace Jovanovich, Inc.
- Arendt, H. (1984). *La vida del espíritu. El pensar, la voluntad y el juicio en la filosofía y en la política*. Centro de estudios constitucionales (Madrid).
- Arendt, H. (1998a). Intercambio epistolar entre Gershom Scholem y Hannah Arendt. *Raíces*, 36, 23–33.
- Arendt, H. (1998b). *Los orígenes del totalitarismo*. Taurus.
- Arendt, H. (2003a). *Eichmann en Jerusalén. Un estudio sobre la banalidad del mal*. Lumen.
- Arendt, H. (2003b). Thinking and Moral Considerations. En Arendt, H., *Responsibility and Judgment* (pp. 159–189). Schocken Books.
- Arendt, H. (2003c). Some Questions of Moral Philosophy. En Arendt, H., *Responsibility and Judgment* (pp. 49–146). Schocken Books.
- Arendt, H. (2006). *Diario filosófico 1950–1973*. Herder.
- Arendt, H. (2007). The Eichmann Controversy. A letter to Gershom Scholem. En Arendt, H., *The Jewish Writings* (pp. 465–471). Schocken Books Press.
- Bernstein, R. (2004). *El mal radical: una indagación filosófica*. Lilmod.
- Blanco Ilari, J. (2013). Comprensión y reconciliación: algunas reflexiones en torno a Hannah Arendt. *Contrastes. Revista Internacional de Filosofía*, XIX(2), 319–338.
- Canclini, R. (2016). Algunas objeciones arendtianas a la analogía orgánica de Rousseau. En Roetti, J. y Moro, R. (Eds.), *Cuadernos de lógica, epistemología y lenguaje. El fundamento y sus límites. Algunos problemas de fundamentación en ciencia y filosofía* (pp. 109–128). College Publications.
- Canovan, M. (1992). *Hannah Arendt. A Reinterpretation of Her Political Thought*. Cambridge University Press.
- Heuer, W. (2005). «La imaginación es el prerequisite del comprender» (Arendt): sobre el puente entre pensamiento y el juzgamiento. *Cuadernos de ética y filosofía política*, 7(2), 37–51.
- Hilb, C. (2015). «Por eso, Sr. Eichmann, debe Ud. colgar». De *Eichmann en Jerusalén* a los «Juicios» en Argentina (reflexiones situadas). *African Yearbook of Rhetoric*, 6(1), 3–13.
- Jerba, H. (2019). Arendt, Scholem y la respuesta del mundo judío al problema del antisemitismo. *Revista de Estudios Políticos*, 183, 13–32.
- Kant, I. (1969). *La religión dentro de los límites de la mera razón*. Alianza Editorial.

- Kateb, G. (2013). Existential Values in Arendt's Treatment of Evil and Morality. En Benhabib, S. (Ed.), *Politics in Dark Times. Encounters with Hannah Arendt* (pp. 342–373). Cambridge University Press.
- Kohn, J. (2003). Introduction. En Arendt, H., *Responsibility and Judgment* (pp. VII–XXIX). Schocken Books.
- Pendas, D. (2007). «Eichmann in Jerusalem», Arendt in Frankfurt: The Eichmann Trial, the Auschwitz Trial, and the Banality of Justice. *New German Critique*, 100, 77–109.
- Serrano de Haro, A. (1998). Acotaciones a la correspondencia. *Raíces*, 36, 34–44.
- Villa, D. (1999). Conscience, the Banality of Evil, and the Idea of a Representative Perpetrator. En Villa, D., *Politics, Philosophy, Terror. Essays on the Thought of Hannah Arendt* (pp. 39–60). Princeton University Press.
- Young–Bruehl, E. (2004). *Hannah Arendt. For Love of the World*. Yale University Press.

## Sobre las autoras y los autores

Nicolás Alles · Licenciado en Filosofía (Universidad Nacional del Litoral). Doctor en Filosofía (Universidad Nacional de La Plata). Docente e investigador (Facultad de Ciencias Jurídicas y Sociales, UNL). Profesor de Ética, Filosofía Política y Filosofía del Derecho en las carreras de Licenciatura y Profesorado en Filosofía (Universidad Autónoma de Entre Ríos). Docente de las asignaturas Filosofía Política Clásica y Filosofía Política Contemporánea en la Maestría en Filosofía (Universidad Nacional de Quilmes).  
nicolas.alles@gmail.com

Federico Arena · Abogado (Universidad Nacional de Córdoba). Doctor europeo en Filosofía del Derecho y Bioética Jurídica (Universidad de Génova). Ha sido investigador posdoctoral (Universidad Bocconi de Milán). Investigador Adjunto (CONICET). Profesor Asociado (Universidad Blas Pascal). Titular interino (Universidad Nacional de Córdoba). Sus líneas de investigación son la Teoría del Derecho y de la interpretación jurídica, por un lado, y la relevancia jurídica de la categorización social (estereotipos, prejuicios y sesgos), por el otro.  
farena@gmail.com

Gustavo Beade · Profesor de Derecho (Facultad de Ciencias Jurídicas y Sociales, Universidad Austral de Chile). Doctor en Derecho (Universidad de Buenos Aires). Áreas de investigación: filosofía moral, filosofía política y filosofía jurídica en especial la filosofía del derecho penal.  
gustavo.bead@uach.cl

Daniel Busdygan · Doctor en Filosofía (Universidad Nacional de La Plata). Dirige la Maestría en Filosofía. Profesor de grado y posgrado (UNLP y Universidad Nacional de Quilmes). Miembro investigador de la Unidad de investigaciones en filosofía legal, jurídica y política (UNQ), Comité de Ética de la investigación (UNQ) y del Centro de investigaciones y estudios de filosofía del Instituto de Investigaciones en Humanidades y Ciencias Sociales (Ciefi-IdiHCS-CONICET). Autor, entre otros, de *Sobre la despenalización del aborto* y, en colaboración, *Cognocimiento, arte y valoración: perspectivas filosóficas actuales, Ideas y perspectivas filosóficas, Aborto: aspecto normativos, jurídicos y discursivos* y *Rostros del igualitarismo*.  
dbusdygan@unq.edu.ar

Rebeca Canclini · Doctora en Filosofía. Profesora de Filosofía política (Universidad Nacional del Sur). Directora del proyecto de investigación «El uso de imágenes para la conceptualización de la tensión unidad/pluralidad en el espacio público: Hannah Arendt y Chantal Mouffe en diálogo con la tradición de pensamiento político». Autora de *Un mundo para la acción: subjetividades políticas y ley en Hannah Arendt*.  
rebeca.canclini@uns.edu.ar

Samuele Chilovi · Graduado en Filosofía (Universidad de Bologna). Doctorado (Universidad de Barcelona, España). Realiza estudios posdoctorales (Universidad de California). Sus áreas de interés son la metafísica y la filosofía del lenguaje aplicadas. Estudia la relación entre las propiedades de validez legal y las cosas más básicas en virtud de las cuales algunas normas poseen tal propiedad. Autor de numerosos artículos en revistas internacionales.  
samuele.chilovi@gmail.com

Martín Daguerre · Profesor y Licenciado en Filosofía (Universidad Nacional de La Plata). Doctor en Sociología (Universidad de Barcelona). Profesor Adjunto de Ética (FAHCE, UNLP), Profesor Adjunto de Lógica (FAHCE y Facultad de Psicología, UNLP), Profesor contratado de Filosofía Política Contemporánea (Universidad Nacional de Quilmes). Dirige el proyecto I+D «La normatividad en ética y en lógica: una perspectiva evolutiva».  
mdaguerre@fahce.unlp.edu.ar

Fernanda Flores · Licenciada en Filosofía (Universidad Nacional del Litoral). Se especializa en temas de filosofía práctica, tales como la motivación y la agencia moral, principalmente desde la perspectiva de la metaética y la ética analítica. Ha participado en proyectos de investigación y de extensión de la UNL y de otras instituciones. Obtuvo becas de grado de iniciación a la investigación (CIN).  
f.flores016@gmail.com

Facundo García Valverde · Doctor en Filosofía y docente (Universidad de Buenos Aires). Investigador en el CONICET. Autor de diversos artículos en *Dianoia*, *Revista de Filosofía*, *Revista Latinoamericana de Filosofía*, *Contrastes* y *Revue de Ethique et Economie*, y de capítulos en compilaciones sobre filosofía práctica. Sus áreas de interés filosófico son el enfoque de las capacidades, concepciones no cognitivistas de la confianza y el análisis normativo de las políticas de protección social.  
fgvalverde@filo.uba.ar

Daniel González Lagier · Doctor en Derecho y catedrático de Filosofía del Derecho (Universidad de Alicante). Ha impartido cursos y conferencias en universidades e instituciones de España, Colombia, Argentina, Italia, México. Autor de *Acción y norma en G.H. von Wright*, *G.H. von Wright y los conceptos básicos del Derecho*, *Las paradojas de la acción*, *Quaestio Facti. Ensayos sobre prueba, causalidad y acción* y de diversos artículos publicados en revistas especializadas.  
daniel.gonzalez@ua.es

Daniel Kalpokas · Doctor en Filosofía (Universidad de Buenos Aires). Investigador Independiente del CONICET. Profesor Adjunto Regular (Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba). Autor de *Richard Rorty y la superación pragmatista de la epistemología* y de varios artículos publicados en revistas nacionales y extranjeras. Áreas de interés: teoría del conocimiento, filosofía de la percepción, filosofía del lenguaje, filosofía de la mente, filosofía contemporánea (principalmente pragmatismo clásico y contemporáneo) y metaética.  
dkalpokas@unc.edu.ar

Guillermo Lariguet · Doctor en Derecho y Ciencias Sociales (Universidad Nacional de Córdoba). Premio Konex al mérito en ética 2016. Investigador Independiente de Conicet. Profesor Asociado de Ética en la carrera de Filosofía (Universidad Nacional del Litoral).  
gclariguet@gmail.com

David Martínez Zorrilla · Licenciado en Derecho (Universitat Autònoma de Barcelona). Doctor *cum laude* en Derecho (Universitat Pompeu Fabra). Profesor Agregado de Filosofía del Derecho en los Estudios de Derecho y Ciencias Políticas (Universitat Oberta de Catalunya). Profesor en ESERP Business School (Barcelona). Autor, entre otros, de *Conflictos constitucionales, ponderación e indeterminación normativa* y *Metodología jurídica y argumentación*.  
dmartinezz@uoc.edu

Guadalupe Mettini · Profesora de Filosofía (Universidad Nacional del Litoral). Doctoranda en Filosofía (Universidad de Buenos Aires). Áreas de investigación: Filosofía de las Ciencias, Epistemología, Enseñanza de la Filosofía.  
guadalupemettini@gmail.com

Lucas Misseri · Profesor y Licenciado en Filosofía (Universidad Nacional de Mar del Plata). Doctor en Filosofía (Universidad Nacional de Lanús). Doctorando en Derecho (Universidad de Alicante), donde es profesor visitante del Departamento de Filosofía del Derecho. Sus principales líneas de investigación tienen dos ejes complementarios: la historia de las ideas (en especial el estudio de las utopías modernas) y la filosofía práctica (teorías éticas, iusfilosóficas y políticas). Sobre esos ejes ha publicado múltiples trabajos académicos, tanto en español como en inglés y eslovaco.

lucas.misseri@ua.es

Martín Oliveira · Doctor en Filosofía (Universidad de Buenos Aires). Se desempeña como profesor Adjunto del Departamento de Ciencia Política (Facultad de Ciencias Sociales, Universidad de Buenos Aires). Jefe de Trabajos Prácticos en el Departamento de Filosofía (Universidad de Buenos Aires). Becario posdoctoral del CONICET. Sus áreas de investigación son el liberalismo igualitario, la justicia distributiva internacional y el problema metodológico en la filosofía política contemporánea.

moliveira@uba.ar

Leandro Paolicchi · Doctor en Filosofía (UNLa, Argentina). Docente (Facultad de Humanidades, Universidad Nacional de Mar del Plata). Investigador Adjunto del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Como becario del Deutscher Akademischer Austauschdienst (DAAD), ha realizado estudios doctorales (2012) y posdoctorales (2014) en la Goethe Universität de Frankfurt am Main bajo la dirección de Axel Honneth. Autor de los libros *Discurso y Facticidad* (Suárez, 2009) y *Praxis, sentido y normatividad. Hacia una reconstrucción pragmática de las acciones humanas* (Prometeo, 2014). Sus principales líneas de investigación son las teorías del discurso y la filosofía social y política contemporánea.

leandropaolicchi@hotmail.com

Luciana B. Samamé · Licenciada y Doctora en Filosofía (Universidad Nacional de Córdoba). Fue becaria posdoctoral CONICET. Ha sido profesora visitante en la Universidad de Granada e investigadora visitante en la Universidad de Murcia. Se ha desempeñado como profesora de Ética y Epistemología (Universidad Blas Pascal, Universidad Provincial de Córdoba y en la Universidad Yachay Tech, Ecuador). Sus inquietudes filosóficas se inscriben en el área de la filosofía práctica y la ética contemporánea, y ha publicado diversos artículos afines a ese

dominio. Actualmente, adscribe sus actividades de investigación al Centro de investigaciones jurídicas y sociales de la UNC.

lucsamame@yahoo.com.ar

Santiago Truccone–Borgogno · Filósofo práctico, trabaja sobre justicia intergeneracional. Sus proyectos actuales versan sobre injusticias históricas, pueblos originarios, cambio climático y transiciones justas. Sus investigaciones pasadas se centraban en la noción de daño, el problema de la no-identidad, y asuntos de derecho penal. Es investigador posdoctoral y predoctoral del departamento de filosofía de la Universidad de Graz, donde también es candidato a PhD en filosofía. Es Doctor en Derecho (Universidad Nacional de Córdoba).

truccone-borgogno@uni-graz.at

Adriana Vercellone · Abogada, Magíster en Derecho y Argumentación. Doctora en Derecho en Ciencias Sociales (Universidad Nacional de Córdoba). Profesora en Ética y en Derecho Político de la Facultad de Derecho (UNC). Becaria Posdoctoral CONICET. Directora del grupo de investigación en «Ética y Función Pública» radicado en el Centro de Investigaciones Jurídicas y Sociales (UNC), subsidiado por la Secretaría de Ciencia y Tecnología (UNC) y el Ministerio de Ciencia de la Provincia de Córdoba. Miembro del Programa de Ética y Teoría Política ([www.programadeetica.com.ar](http://www.programadeetica.com.ar)).

adrivercellone@gmail.com

Graciela Vidiella · Doctora en Filosofía (Universidad de Buenos Aires). Profesora Extraordinaria en la categoría Consulto (Universidad Nacional de La Plata). Áreas de investigación: teorías de la democracia, racionalidad práctica y emociones; ética y literatura.

gravidiella@gmail.com

María E. Wagon · Licenciada y Doctora en Filosofía (Universidad Nacional del Sur) y becaria posdoctoral del CONICET. Docente ayudante en la cátedra de Filosofía de la Religión del Profesorado y la Licenciatura en Filosofía (UNS). Intereses de investigación: dentro de la teoría y filosofía política, temas relacionados con el pensamiento de Hannah Arendt y las cuestiones de género. Miembro del Centro Interdisciplinario de Estudios de Género y Feminismos (CIEGEF, UNS).

mariawagon@gmail.com



María Sol Yuan · Doctora en Filosofía (Universidad de Buenos Aires). Investigadora Asistente de CONICET. Profesora Adjunta en Teoría del Conocimiento (Facultad de Humanidades y Ciencias, Universidad Nacional del Litoral). Áreas de interés: teorías del conocimiento, problemas en torno a la percepción, filosofía de Wittgenstein, epistemología social y de género. Pragmatismo.  
myuan@unl.edu.ar

La ética es una de las disciplinas más respetadas y desarrolladas de toda la tradición filosófica. Casi no ha habido autor de importancia que no haya dedicado parte de sus esfuerzos a pensar distintos aspectos de esta rama de la filosofía. La posibilidad de concebir hechos morales objetivos, la incidencia de las perspectivas naturalistas en el razonamiento moral, la dimensión metafísica de la moral son solo algunos de los tantos temas que demandan un ámbito de reflexión específico en el que convergen, junto con la ética, áreas de la filosofía tales como la filosofía del lenguaje, la metafísica, la epistemología, e incluso la psicología moral. Estos cruces constituyen el ámbito propio de la metaética, rama de la ética contemporánea que ha experimentado un gran crecimiento en los últimos años y que convoca la atención de algunos de los más relevantes especialistas de la filosofía moral de la actualidad.

Aunque importante en el panorama de las discusiones filosóficas, la metaética todavía no ha tenido, en nuestro medio, la atención que merece.

*La metaética puesta a punto* pretende contribuir a saldar esta deuda al poner a disposición del público de habla hispana el tratamiento de algunos de los temas de esta área del pensamiento ético. Los capítulos que lo componen se dirigen a estudiantes e investigadores en el área de las humanidades interesados en la discusión moral y los cruces que pueden surgir a partir de ella.