



La filosofía como una disciplina humanística

Bernard Williams

Selección, coordinación e introducción

A. W. Moore



La filosofía como una disciplina humanística

Bernard Williams

Traducción *Adolfo García de la Sienna*
Selección, coordinación e introducción *A. W. Moore*



Primera edición, 2011
Primera edición electrónica, 2011

Título original: *Philosophy as an Humanistic Discipline*

D. R. © 2011, Fondo de Cultura Económica
Carretera Picacho-Ajusco, 227; 14738 México, D. F.
Empresa certificada ISO 9001:2008



www.fondodeculturaeconomica.com

Comentarios:
editorial@fondodeculturaeconomica.com
Tel. (55) 5227-4672
Fax (55) 5227-4649

Se prohíbe la reproducción total o parcial de esta obra, sea cual fuere el medio. Todos los contenidos que se incluyen tales como características tipográficas y de diagramación, textos, gráficos, logotipos, iconos, imágenes, etc. son propiedad exclusiva del Fondo de Cultura Económica y están protegidos por las leyes mexicana e internacionales del copyright o derecho de autor.

ISBN 978-607-16-0805-5

Hecho en México - *Made in Mexico*

Acerca del autor

Bernard Williams fue investigador en el All Souls College de Oxford. Fue profesor en la Universidad de California en Berkeley desde 1988 hasta su muerte, y en Oxford desde 1990 hasta 1996. Sus libros incluyen *Truth and Truthfulness* (Princeton), *Morality, Descartes*, *Ethics and the Limits of Philosophy*, y *Shame and Necessity*.

Índice

Prefacio. Patricia Williams

Introducción. A. W. Moore

Primera parte. Metafísica y epistemología

- I. La paradoja de Tertuliano
- II. Argumentos metafísicos
- III. Placer y creencia
- IV. Conocimiento y razones
- V. Identidad e identidades

Segunda parte. Ética

- VI. La primacía de las disposiciones
- VII. La estructura de la teoría de Hare
- VIII. Subjetivismo y tolerancia
- IX. El *actus reus* del doctor Caligari
- X. Los valores, las razones y la teoría de la persuasión
- XI. Responsabilidad moral y libertad política
- XII. Tolerar lo intolerable
- XIII. El prejuicio humano

Tercera parte. El alcance y los límites de la filosofía

- XIV. La filosofía política y la tradición analítica
- XV. La filosofía y el entendimiento de la ignorancia
- XVI. La filosofía como una disciplina humanística
- XVII. ¿En qué podría convertirse la filosofía?

Bernard Williams: Publicaciones filosóficas completas

Prefacio

Patricia Williams

Es triste, si bien apropiado, que mi gesto final, práctico, de aprecio y amor para Bernard debiera ser para ayudar a la publicación de las últimas tres colecciones de sus escritos filosóficos. *El sentido del pasado: ensayos en la historia de la filosofía*, *La filosofía como disciplina humanística*, y *En el principio era el hecho: realismo y moralismo en el argumento político*, serían publicados por la Princeton University Press tres años después de su muerte en junio de 2003. Bernard me ayudó y me animó de maneras innumerables en mi carrera de editora, confirmando mi convicción de que los editores de las prensas universitarias debieran ser juzgados por su elección de asesores tanto como por los autores que publican.

Como muchos que lo conocieron, pensé que Bernard era indestructible, y pienso que ¡él pensaba lo mismo! Pero cuando se recuperaba de los drásticos efectos de su primera lucha —tratamiento— contra el cáncer en 1999 hablamos por primera, y quizá única vez, acerca de lo que debiera ocurrir a sus artículos si no pudiese terminar *Truth and Truthfulness*. Afortunadamente, lo publicó en 2002, aunque lo hubiera expandido de varios modos si el tiempo no hubiera sido tan apremiante. Lo que aprendí de esta conversación fue que Bernard no tenía fe en su habilidad, ni en la de ningún filósofo, para predecir cuáles de sus trabajos serían de interés duradero para sus sucesores. Eso es algo que el futuro tendría que decidir. Así que, aunque estaba totalmente en contra de lo que llamaba “listas de lavandería” póstumas, se rehusó a expresar cualquier otra opinión acerca de lo que debiera publicarse después de su muerte. Afortunadamente para mí, especificó que, aunque yo debía manejar los aspectos prácticos de la publicación como me pareciera adecuado, él solicitaría “un filósofo joven de resuelta integridad y severidad de juicio que entendiese lo que él había estado tratando de hacer en filosofía” para asegurar que no me apartase del buen camino filosófico. Ese filósofo fue Adrian Moore. Le estoy profundamente agradecida por la cuidadosa consideración que ha tenido con los complicados asuntos de la publicación y la republicación, y por su amistad. Él es el único arquitecto de este volumen particular.

Mi agradecimiento de corazón, también, a Walter Lippincott, el director de la Princeton University Press, y a su equipo en Princeton y Oxford, cuyo compromiso con Bernard como autor y con los altos estándares de edición, diseño, producción y mercadeo es tan valioso en un tiempo en que la publicación académica encara complejos desafíos financieros.

Finalmente, me gustaría reconocer a los editores que amablemente han otorgado su permiso para publicar material en este volumen.

- I. “La paradoja de Tertuliano”, reimpresso con autorización de Scribner, una edición de Simon & Schuster Adult Publishing Group, de *New Essays in Philosophical Theology*, coordinado por Anthony Flew y Alasdair MacIntyre. Copyright © 1995 de Anthony Flew y Alasdair MacIntyre.
- II. “Argumentos metafísicos”, en D. F. Pears (coord.), *The Nature of Metaphysics*, Macmillan, Londres, 1957. Reproducido con autorización de Palgrave Macmillan.
- III. “Placer y creencia”, *Proceedings of the Aristotelian Society* (suplemento), vol. 33, Blackwell, Oxford, 1959.
- IV. “Conocimiento y razones”, en G. H. von Wright (coord.), *Problems in the Theory of Knowledge*, Nijhoff, La Haya, 1972. © 1972 por Martinus Nijhoff, La Haya, Países Bajos, con la amable autorización de Springer Science y Business Media.
- V. “Identidad e identidades”, en Henry Harris (coord.), *Identity*, Oxford University Press, Oxford, 1995, con autorización de Oxford University Press.
- VI. “La primacía de las disposiciones”, en Graham Haydon (coord.), *Education and Values: The Richard Peters Lectures*, Instituto de Educación de la Universidad de Londres, Londres, 1987.
- VII. “La estructura de la teoría de Hare”, en Douglas Seanor y N. Fotion (coords.), *Hare and Critics: Essays in Moral Thinking*, Oxford University Press, Oxford, 1988, con autorización de Oxford University Press.
- VIII. “Subjetivismo y tolerancia”, en A. Phillips Griffiths (coord.), *A. J. Ayer: Memorial Essays*, Cambridge University Press, Cambridge, 1992.
- IX. “El *Actus Reus* del doctor Caligari”, *University of Pennsylvania Law Review* 142, 1994.
- X. “Los valores, las razones y la teoría de la persuasión”, en Francesco Farina, Frank Hahn y Stefano Vannucci (coords.), *Ethics, Rationality, and Economic Behaviour*, Oxford University Press, Oxford, 1996, con autorización de Oxford University Press.
- XI. “Responsabilidad moral y libertad política”, *Cambridge Law Journal* 56, Cambridge University Press, Cambridge, 1997.
- XII. “Tolerar lo intolerable”, en Susan Mendus (coord.), *The Politics of Toleration in Modern Life*, Duke University Press, Durham, 2000.
- XIV. “La filosofía política y la tradición analítica”, en Melvin Richter (coord.), *Political Theory and Political Education*, Princeton University Press, Princeton, 1980.
- XV. “La filosofía y el entendimiento de la ignorancia”, *Diogène* 169 (1995). También publicado en inglés por Berghahn Books, Oxford.
- XVI. “La filosofía como una disciplina humanística”, *Philosophy* 75, Cambridge University Press, Cambridge, 2000.

Introducción

A. W. Moore

Bernard Williams (1929-2003) fue uno de los más grandes filósofos británicos del siglo XX. Su trabajo, que fue inusual en su alcance, siempre estuvo marcado por una igualmente inusual combinación de rigor, imaginación y profundidad, así como por una plena humanidad. Los ensayos publicados aquí dan copioso testimonio de estas y otras facetas de su extraordinario intelecto, y engloban en cierto modo la entera carrera de Williams: el ensayo I, “La paradoja de Tertuliano”, fue su primera publicación. Apareció hace medio siglo. El ensayo XIII, “El prejuicio humano”, fue dado como conferencia poco antes de que muriese.

Williams mismo produjo tres colecciones de ensayos durante su vida: *Problems of the Self*,^[1] *Moral Luck*,^[2] y *Making Sense of Humanity*.^[3] Preparó una cuarta, dedicada a la historia de la filosofía, que fue publicada póstumamente junto con algunas piezas sobre Nietzsche que pretendía fuesen parte de otro trabajo, con el título de *The Sense of the Past*.^[4] Una quinta colección, de ensayos políticos, en buena medida no publicados hasta ahora, que juntos esbozan a grandes rasgos un libro proyectado sobre política para el cual dejó un bosquejo, se publicó también póstumamente bajo el título de *In the Beginning was the Deed*.^[5] La presente colección, que no se traslapa con ninguna de éstas, completa el conjunto.

No contiene, sin embargo, todo lo que se le hubiera podido añadir. Sigue habiendo ensayos de Williams, tanto publicados como no publicados, que nunca han sido puestos en una antología, y que no he incluido aquí o bien porque se traslapan con otro trabajo suyo publicado, o bien porque eran demasiado ocasionales. (Elaboro más sobre esto en una nota al final de esta introducción.)

En los prefacios a sus propias colecciones, Williams aduce bases similares para excluir material, pero también da razones temáticas. En el prefacio a *Problemas del yo*, escribe: “he dejado algunos artículos fuera debido al tema del que hablan (todo lo que se halla aquí se relaciona sólo con dos o tres temas)”. Esto significa que una de las virtudes de la presente colección es que provee una bienvenida oportunidad para reimprimir los primeros ensayos de Williams, sobre temas acerca de los cuales, por lo demás, escribió muy poco; ensayos en los que encontramos algunos de los más finos ejemplos de su destreza analítica y su claridad de visión.

No es que la presente colección sea una mezcolanza. Hay una unidad de intereses, que se mantiene a lo largo de la carrera de Williams, que evita ese resultado. Una de las

razones por las que me he apropiado el título del ensayo XVI para la colección, es que éste expresa con precisión dicha unidad de intereses. El modo de filosofar de Williams es siempre profundamente autoconsciente dado que, incluso cuando no está explícitamente reflexionando sobre el carácter de la disciplina, su trabajo muestra un agudo sentido tanto de sus posibilidades como de sus limitaciones; del modo de explotar su potencial en tanto disciplina humanística, como de la forma de refrenar sus pretensiones de ser otra cosa.

Para dar una mejor idea de lo que quiero decir con esto, empezaré comentando algo acerca de los tres grupos en que los ensayos están divididos. Al final resistí la tentación de dar a estos grupos las etiquetas “filosofía dura”, “filosofía suave” y “metafilosofía” — aunque pienso que la misma absurdidad de estas etiquetas, combinada con el hecho de que transmiten, en su propio y rudo modo, lo que se busca, hubiera sido atractiva para Williams—. Una función que ciertamente hubieran tenido estas etiquetas habría sido resaltar cómo los dos primeros grupos trazan el terreno de la investigación filosófica de nivel básico, mientras que el tercero provee una visión a ojo de pájaro de ese terreno. Pero hay muchos peligros en los que dichas etiquetas hubieran incurrido, de los cuales el de depreciar al segundo grupo es solamente el más obvio. Otro —o más bien, un peligro en que el agrupamiento ya incurre y que las etiquetas hubiesen exacerbado— es el de sugerir que las divisiones entre los grupos son mucho más nítidas que en la realidad. Varios de los ensayos pudieron haber encajado en diferentes grupos. Por esa razón, entre otras, jugué con la idea de prescindir enteramente del agrupamiento y presentar los ensayos de manera puramente cronológica. Lo que es impactante es cuán pequeña diferencia hubiera hecho esto. La lista como se halla se encuentra a unos pocos ajustes menores de ser puramente cronológica. Esto me parece importante.

Conforme Williams avanzaba de la empresa “más dura” de tratar de encontrar el sentido de nuestro pensamiento y experiencia en general, a la “más suave” de tratar de encontrar el sentido de nuestro pensamiento y experiencia éticos, se volvió cada vez más autoconsciente de la pretensión que tenía la filosofía de ser digna de tomarse con seriedad cuando la magnitud de sus preguntas no era meramente la de la pura generalidad, característica de la metafísica, sino la magnitud, que es (o debería ser) característica de la ética: la magnitud de la *importancia*.^[6] Williams nunca estuvo dispuesto simplemente a dar por sentada la supremacía del argumento racional abstracto para las cuestiones importantes. En *Ethics and the Limits of Philosophy*, cuyo mismo título, por supuesto, es muy sugerente, empezó con lo que tal vez sea la pregunta más importante de todas: cómo debería uno vivir, y dejó en claro desde el principio cuán poco podemos esperar de la filosofía con respecto a esta pregunta: ayuda para entenderla, quizá; ciertamente no para darle una respuesta. (Ésta es una de las ironías de la filosofía moral de Williams, que mientras que sirve como una gloriosa ilustración de cuánto puede lograr la filosofía moral, está dedicada en buena medida a determinar cuán poco la filosofía moral puede lograr.) No es de admirar, entonces, que la autoconciencia que siempre se había hallado al menos implícita en el trabajo de Williams se haya ido haciendo más y más explícita conforme sus intereses filosóficos se volvieron más y más concretos, hasta que finalmente una de sus preocupaciones principales no fue ninguna de

la grandes cuestiones “duras” de la filosofía, ni tampoco sus grandes preguntas “suaves”, sino la naturaleza y prospectos de la filosofía misma.

Debo decir de inmediato que, aunque este modo de caracterizar la evolución del trabajo de Williams tiene casi algo de hegeliano —con su sugerencia de un crecimiento de la autoconciencia realizado a través de una preocupación cada vez más concreta por las realidades de la experiencia ética—, cualquier semejanza con el espíritu del mundo hegeliano es una mera coincidencia. Ciertamente, una cosa está clara: no importa cuán ofendido, incómodo o divertido hubiese estado Williams al ser asociado con Hegel de este modo, se hubiera resistido enérgicamente a cualquier sugerencia de que su obra fuera un esfuerzo por lograr algo que ameritara el título de “conocimiento absoluto”.

No es que negase la posibilidad de tal cosa. Por el contrario, una de las mejor conocidas y más ferozmente impugnadas de sus concepciones filosóficas es la posibilidad de que algo amerite el título de “conocimiento absoluto”, a saber, un conocimiento que sea “en la mayor medida posible independiente de las perspectivas o idiosincrasias locales de los investigadores”.^[7] (Esta definición, aunque menos exigente que otras que se han aventurado, es aun muy exigente. Excluye, por ejemplo, cualquier conocimiento que dependa en algún modo esencial del aparato sensorial que es común en ciertos investigadores, digamos nuestro conocimiento de que el pasto es verde. Nuestro concepto de verdor, vinculado como lo está a nuestro aparato visual, “no estaría disponible a todo observador competente del mundo”).^[8] Pero aunque Williams creía que tal conocimiento absoluto es posible, también creía que, si ha de ser adquirido en alguna parte, ésta es la ciencia, no la filosofía. Esa es una de las principales tesis del ensayo que da nombre a esta colección.

Este extraordinario trabajo es una especie de manifiesto de la concepción de Williams del trabajo de su propia vida. Es aquí donde aborda, de la manera más explícita, hasta dónde la filosofía puede y no puede contribuir al proyecto de encontrar el sentido de las cosas. Niega que la filosofía pueda resultar en conocimiento absoluto porque niega incluso que ello sea un objetivo de la filosofía —como sí lo es, desde su perspectiva, un objetivo de la ciencia—. Pero, insiste, los objetivos de la filosofía son simplemente diferentes, no inferiores. Hábilmente identifica el cientificismo latente de quienes, en una exhibición de anticientificismo, niegan que incluso la ciencia pueda resultar en conocimiento absoluto, pero que lo hacen con desilusión, o con un sentido de alivio concerniente a cualquier comparación potencial con sus propias tareas no científicas o, desde luego, en desafiante reacción a la arrogancia de la ciencia, exhibiendo así su convicción de que, si la ciencia *pudiese* resultar en conocimiento absoluto, entonces sería el Santo Grial de cualquier actividad intelectual. De cada una, Williams demanda lo que le corresponde. En lo que la filosofía puede contribuir (de manera más ideal) al proyecto de encontrar el sentido de las cosas es en encontrar el sentido de *ser humano*, y no se da un gran servicio a dicha contribución abstrayendo “las perspectivas o idiosincrasias locales” de los seres humanos.

Estas ideas acerca del conocimiento absoluto son uno de los legados más importantes de Williams, pero han sido muy mal entendidos y mal representados por innumerables

críticos. En este trabajo (el ensayo XVI), Williams descose algunos de dichos malos entendidos. En particular, desmiente la confusión sorprendentemente común de que, al ser un paladín de la idea del conocimiento absoluto, o de la idea de un saber absoluto del mundo, como a veces lo decía, se volvía un paladín de la idea de una concepción sin conceptos. (Nunca abogó por un absurdo tal, pero frecuentemente se cree que es eso a lo que llegan sus planteamientos.) Una fuente de este malentendido es la idea de que cualquier conjunto de conceptos debe implicar sus propios principios distintivos de asimilación y discriminación, los cuales deben reflejar ciertas preocupaciones e intereses, que a su vez deben depender, en buena medida, de ciertas perspectivas e idiosincrasias locales, pero es precisamente este último paso el que Williams repudia. Un saber absoluto no es una concepción sin conceptos. Es una concepción con conceptos de un tipo especial.

Un malentendido algo más sutil es que Williams defiende la idea de una concepción con conceptos por cuyos medios, y solamente por cuyos medios, “los hechos pueden imprimirse directamente en nuestras mentes, sin ninguna necesidad de mediación por nada tan históricamente condicionado y abierto a disputa como los cánones del argumento científico bueno y malo”.^[9] Nuevamente se piensa que cualquier mediación implica perspectivas e idiosincrasias locales. Pero, ¿por qué? A menos que “directamente” se interprete de modo que deje sin resolver toda suerte de preguntas, no hay nada en Williams que sugiera que los hechos se impriman en nuestras mentes de una manera más directa cuando los concebimos en términos absolutos que cuando los concebimos de cualquier otro modo. Los conceptos incluidos en un saber absoluto siguen siendo conceptos, y median en cualquiera que sea el modo en que los conceptos median entre los hechos y nuestras mentes; mientras el conocimiento que no es absoluto siga siendo conocimiento, y cualquier conocimiento, o al menos cualquier conocimiento proposicional, se puede decir, valga la redundancia, que es conocimiento de los hechos.

Este último punto merece una elucidación. El conocimiento proposicional, tomando prestada la propia definición de Williams en el ensayo IV, “Conocimiento y razones”, es “conocimiento cuya expresión paradigmática en los usuarios del lenguaje es la confiada aseveración de verdades, y donde la pretensión de que es conocimiento lo que está siendo expresado implica como condición necesaria que lo que se afirma sea verdadero”.^[10] Esto excluye los conocimientos “prácticos”, como mi conocimiento del modo de atar mis agujetas, cuya expresión paradigmática es mi acción efectiva de atarlas. Por otra parte, incluye mucho conocimiento que no es absoluto, tal como (para regresar a un ejemplo anterior) mi conocimiento de que el pasto es verde. Y a menos que “hecho” se entienda en algún modo especialmente ambicioso, las “verdades” en la definición de Williams pueden ser sustituidas por “hechos”, y “ser verdadero” por “es un hecho”, lo cual es tanto como decir que cualquier conocimiento proposicional es conocimiento de algún hecho.

Williams, entonces, no dice nada para sugerir que el conocimiento absoluto implique extrañamente un acceso a los hechos sin mediación. En particular, no dice nada para sugerir que tal conocimiento pueda ser alcanzado sin mediación por cánones de

argumento científico bueno y malo. Es cierto que los cánones del argumento científico bueno y malo están “históricamente condicionados”, así como lo están otras partes de nuestra vida intelectual. Esto es algo en lo que Williams sería el primero en insistir, como desde luego lo hace en el ensayo XVI. Nos recuerda que los conceptos científicos tienen una historia; el punto, sin embargo, es que dicha historia es “parte de la historia del descubrimiento”.^[11] El avance de un conjunto de conceptos científicos a otro, o de un canon de argumento científico bueno y malo a otro, puede ser visto desde ambas perspectivas como simplemente eso: un avance. Es una mejoría.

Este es uno de los muchos puntos en que los conceptos científicos difieren de los que usamos frecuentemente para registrar lo que sucede alrededor de nosotros: lo que Williams de manera adecuada ha llamado conceptos éticos “densos”.^[12] La idea de los conceptos éticos densos es otro de los principales legados de Williams. Es de gran importancia para él: como lo testifica el hecho de que aparece en no menos de siete de los ensayos de esta colección (aunque no siempre con esa etiqueta). Por un concepto ético denso Williams entiende un concepto que (a diferencia de un concepto científico) tiene un aspecto evaluativo —pero que también (a diferencia de un concepto ético “ligero”, tal como el de hacer el mal, que en la perspectiva de Williams no puede servir, de ninguna manera directa, para “reportar” nada) tiene un aspecto fáctico. Así que, aplicar un concepto ético denso en una situación dada es, en parte, evaluar la situación, pero es también decir algo directamente falso si la situación no resulta ser de un modo determinado. Un ejemplo es el concepto de infidelidad. Si te acuso de ser infiel, con ello te censuro, pero digo también algo de lo que estoy obligado a retractarme si resulta que, de hecho, no has traicionado ningún acuerdo relevante. Otros ejemplos son los conceptos de blasfemia, castidad, valentía y pereza.

Las diferencias que Williams reconoce entre conceptos científicos y conceptos éticos densos se reflejan en la distinción que hace entre las creencias científicas y las creencias éticas. Sostiene que las primeras gozan de una especie de objetividad de la que carecen las segundas. Esto está conectado con la posibilidad de que alcancemos un acuerdo diáfano sobre asuntos científicos o, como a veces lo dice Williams, de que *converjamos* en nuestras creencias científicas,^[13] como lo opuesto a la posibilidad de alcanzar un acuerdo diáfano sobre asuntos éticos o *converjamos* en nuestras creencias éticas. Pero la concepción no es que, mientras que podemos razonablemente esperar hacer lo primero, no podamos razonablemente esperar hacer lo segundo. Mucho menos es que a veces hagamos lo primero pero nunca lo segundo. Ni tiene que ver con que, donde haya *convergencia*, las creencias en cuestión constituyan, o no, conocimiento. Tiene que ver con los diferentes modos de explicar cualquier cosa que sea la convergencia.

La concepción de Williams es como sigue: Las personas a veces convergen en sus creencias éticas, y aquellas creencias a veces constituyen conocimiento. Esto puede suceder precisamente cuando las creencias en cuestión implican un concepto ético denso. Así, las personas que abrazan el concepto de blasfemia podrían no tener dificultad para estar de acuerdo ni, desde luego, para saber que cierta obra de arte, digamos, es blasfema. Sin embargo, lo central es, en primer lugar, lo que tenga que ver con abrazar el

concepto de blasfemia. Concedida la distintiva combinación de evaluación y objetividad del concepto, abrazarlo es parte de vivir en un mundo social particular, un mundo en el que ciertas cosas son apreciadas y otras aborrecidas. Las personas necesitan vivir en tal mundo social. Pero como la historia lo demuestra ampliamente, no hay tal mundo social en que las personas necesiten vivir. Desde luego no necesitan vivir en un mundo que sustente el concepto de blasfemia, así que cualquier buena explicación reflexiva acerca de por qué las personas convergen en sus creencias acerca de lo que es blasfemo, debe incluir una explicación científico-social de por qué abrazan finalmente el concepto de blasfemia; por qué viven en *ese* mundo social. Esta explicación no puede por sí misma invocar el concepto de blasfemia, porque debe hacerse desde un punto de vista privilegiado; fuera del mundo social en cuestión. Así que no puede conformarse al esquema “Esas personas convergen en sus creencias acerca de x porque son, de manera idónea, sensibles a las verdades acerca de x ”. Esto es, no se les puede representar en consonancia sobre lo que es blasfemo sólo por la perspicacia que tienen acerca de lo que es blasfemo. En contraste, una buena explicación reflexiva de por qué las personas convergen en sus creencias acerca de (digamos) cómo es el nitrógeno, por tomar un ejemplo científico estándar, *puede* por sí mismo invocar el concepto de nitrógeno y, por ende, suponiendo que las creencias han sido alcanzadas propiamente, *pueden* conformarse al esquema especificado antes. Se puede representar a esas personas en consonancia acerca de cómo es el nitrógeno debido a que lograron un mejor entendimiento acerca de cómo es el nitrógeno; debido a lo que han descubierto acerca del nitrógeno.

Esto es, desde luego, una variación del tema de que el conocimiento ético es dependiente de “las perspectivas o idiosincrasias locales de los investigadores”, mientras que el conocimiento científico no puede serlo. Qué conceptos éticos abrazan las personas es desde luego parte de lo que determina sus perspectivas e idiosincrasias locales. Una buena explicación reflexiva de cómo las personas tienen el conocimiento ético que tienen debe, por lo tanto, incluir una explicación de cómo tienen algunas de las perspectivas e idiosincrasias locales que tienen, y no puede hacer esto a menos que se separe de aquellas perspectivas e idiosincrasias. En contraste, una buena explicación reflexiva de cómo las personas tienen el conocimiento científico que tienen, que puede ser absoluto, no requiere seguir el mismo tipo de rodeo.

Estas reflexiones sobre la reflexión nos retrotraen a la cuestión de la filosofía. La filosofía, obviamente, implica reflexión. Y la reflexión, a su vez, implica objetividad. Qué grado de objetividad es apropiado en cualquier ejercicio dado de reflexión, depende de la meta del ejercicio. (En el caso que hemos estado considerando, donde la meta es explicativa y donde varias perspectivas e idiosincrasias locales son en sí mismas los *explicanda*, el grado de objetividad requerido, al menos de aquellas perspectivas e idiosincrasias locales, es total). Muy bien, ¿qué grado de objetividad es adecuado a cualquiera de los múltiples objetos de la filosofía? No, dado lo que se ha dicho antes, el grado absoluto al que los científicos podrían aspirar. Pero suficiente, en muchos casos, para que nuestros conceptos éticos densos dejen de estar entre las cosas con las que

pensamos y lleguen a estar entre las cosas acerca de las cuales pensamos como, desde luego, se acaba de hacer. El proyecto de encontrar el sentido del ser humano requiere que se reconozca que nuestros conceptos éticos densos son fenómenos contingentes, cuyas historias no suelen reivindicarlos, cuyas contribuciones a nuestras vidas se hallan continuamente modificadas por toda suerte de fuerzas sociales movedizas, y cuyo futuro mismo es cuestionarse^[14] —y, aunque puedan ser capaces de sustentar cierto tipo de objetividad en nuestro pensamiento ético, son incapaces, por las razones que acabamos de considerar, de apoyar otros.

El problema, como lo destaca Williams en el ensayo XVI, es que esto puede ser muy perturbador. Hacerse conscientes de la fragilidad de nuestros conceptos éticos densos, y de la existencia de alternativas, puede reducir la firmeza con que los asimos. Más aún, puesto que algunos objetivos filosóficos exigen menos objetividad que esto, desde luego lo suficientemente poco como para que sigamos asiendo nuestros conceptos éticos densos y para pensar crítica e imaginativamente *con* ellos, se deduce que hay una cierta tensión dentro de la misma filosofía. A ésta regresa constantemente el trabajo de Williams, y está relacionada con una de las tensiones a las que el argumento de su libro *Ethics and the Limits of Philosophy* vuelve una y otra vez, de acuerdo con lo que dice en el *postscript*: la tensión entre reflexión y práctica.^[15] Con el desapego que hace posible que nuestro pensamiento y práctica éticos se vuelvan objetos de investigación, contiene el acuerdo que hace posible que tengamos un pensamiento y una práctica absolutamente éticas, un acuerdo basado en lo que Williams ha identificado de varios modos como compromiso,^[16] convicción,^[17] y confianza.^[18]

Es un lugar común que la autoconciencia y la autoconfianza no van fácilmente juntas. La perspicacia de Williams nos lleva más allá de ese lugar común de varios modos. Uno de éstos consiste en ubicar la misma relación incómoda en el dominio de lo social. Otro es ubicarla en el dominio del pensamiento, específicamente del pensamiento ético. Y un tercero es ubicarla en el contexto de un pluralismo de valores. Pues aquí hay, en la concepción de Williams, bienes en competencia. La reflexión es un bien, como se ha dicho de manera adecuada, si bien quizá magnificado, en el dicho socrático de que no vale la pena vivir una vida que no ha sido examinada.^[19] Y la confianza es un bien, sin el cual no habría ninguna cosa como vivir una vida, o no en un sentido más allá del meramente biológico. Pero ninguno de estos bienes es un bien supremo. Cada uno consume al otro; cada uno “tiene un precio, que no debería fijarse demasiado alto”.^[20]

Cómo equilibrar estos bienes es en sí misma una cuestión práctica, y abordarla propiamente requiere un equilibrio de reflexión y confianza. En buena medida debido a la conciencia que Williams tiene de esto, su obra filosófica siempre ha tenido un agudo sentido de sus posibilidades y limitaciones; de sus beneficios y sus peligros potenciales. Esto es especialmente claro en el ensayo XVI, y en los otros ensayos del tercer grupo, donde lo que está haciendo Williams es precisamente reflexionar sobre la filosofía. Pero hay evidencia de ello a lo largo de la colección en la que Williams es también, desde luego, un practicante, cuyo trabajo está marcado no por una indiscriminada confianza en la filosofía, sino por una confianza en lo que él mismo puede producir con éxito: la

filosofía en su mejor forma.[21]

Nota sobre la selección

Hay ensayos de Williams, tanto publicados como no publicados, que no incluyo en esta colección aunque no aparecen en ninguna de sus otras colecciones. Esto lo hago por diversas razones: *a*) traslape con los ensayos que ya he incluido (por ejemplo, he excluido varios ensayos sobre tolerancia, tema que aparece *verbatim* en el ensayo XII, “Tolerar lo intolerable”); *b*) traslape con, o desplazamiento por otro trabajo publicado por Williams (por ejemplo, he excluido varios ensayos que son en efecto primeros bosquejos de capítulos de *Verdad y veracidad*,[22] o cuyas ideas principales han sido incorporadas en ese libro), y *c*) por ser demasiado ocasionales (por ejemplo, he excluido varios ensayos que son respuestas directas al trabajo de otras personas y que aislados tendrían muy poco sentido).

Al final del libro aparece una lista completa de los ensayos publicados de Williams, incluyendo los que he dejado fuera. (Esta bibliografía también contiene una lista de las reseñas de Williams, sólo unas pocas de las cuales han sido puestas en una antología hasta el momento. Un volumen separado de éstas podría aparecer en una fecha posterior.)

Todos los ensayos que aparecen aquí, a excepción de dos, han sido publicados con anterioridad (aunque muchos de ellos eran prácticamente inaccesibles).[23] Las dos excepciones son el ensayo XIII, “El prejuicio humano”, y el XVII, “¿En qué podría convertirse la filosofía?”, que fue presentado como conferencia inaugural para el Centro para la Filosofía Postanalítica en la Universidad de Southampton.[24] Los orígenes de los ensayos restantes quedaron registrados en la sección de reconocimientos al final del prefacio.

Primera Parte

Metafísica y epistemología

I. La paradoja de Tertuliano[a]

*Non pudet, quia pudendum est... prorsus credibile est, quia ineptum est...
certum est, quia impossibile.*
Tertuliano, *De carne Christi*, v.

1

Este artículo no trata directamente ni de Tertuliano ni de su paradoja. Al considerar la más famosa y más ampliamente citada —incorrectamente— de las paradojas de Tertuliano, no trato de explicarla, mucho menos de disolverla; la tomo como el punto de partida y fin de una discusión del lenguaje religioso y de sus relaciones con la teología y con el tipo de indagación filosófica de la que este libro se ocupa principalmente.[b] En particular, trato de hacer manifiesta una cierta tensión, un estira y afloja entre lo posible y lo imposible; una especie de incomprendibilidad inherente y necesaria que parece ser una característica de la creencia cristiana, y ubicar este punto de tensión más exactamente dentro de la estructura de la creencia. Tertuliano parece haber sentido muy fuertemente esta tensión y de forma original la proclamó con vigor, pero es sólo por este cordel más bien estrecho que mis observaciones están atadas a lo que dijo Tertuliano, cuya estricta interpretación requeriría algo muy diferente.

Como la trayectoria de este artículo es más bien tortuosa, puede ayudar un tosco mapa. Después de formular la paradoja (2), procedo a una breve discusión de las paradojas en general, sus usos y exigencias (3). Dejo entonces a Tertuliano por un momento e intento mostrar algunas características que distinguen el lenguaje religioso, cuando menos el cristiano, de otros tipos de lenguaje (4); esto se hace presuponiendo la existencia de Dios, lo cual puede parecer un procedimiento más bien peculiar a un escéptico pero que servirá, espero, para una discusión que trata de mostrar algo acerca del lenguaje religioso tal y como lo usan los creyentes. Se propone entonces la tesis de que la creencia cristiana debe implicar al menos una proposición que hable tanto de Dios como del mundo, y que dicha proposición debe ser parcialmente incomprendible —lo cual sostengo que es sugerido por la paradoja de Tertuliano, si ha de tomarse al pie de la letra (5). Se hacen entonces algunas observaciones sobre la teología y sus relaciones con el lenguaje religioso y con la filosofía del lenguaje religioso; éstas dan lugar a consideraciones que ponen un alto a una incipiente discusión sobre la Encarnación y

sugieren algunas conclusiones más bien descorazonadoras acerca tanto de la filosofía del lenguaje religioso como de la teología (6). Termino (7) con algunas observaciones acerca de la fe y de lo que uno puede o no puede decir que cree por fe.

De acuerdo con mi representación, la paradoja de Tertuliano habla tanto de la creencia cristiana como de la teología, pero lo primero es lo más importante. En ambos casos la considero como una paradoja acerca del significado más que acerca de la verdad; esto es, me ocupo de lo que se dice en el lenguaje religioso, y no tanto de averiguar si es verdadero o falso, aunque las dos cuestiones no están (ni pueden estar) fríamente separadas.

2

Tertuliano, el primer padre latino de la Iglesia, empezó su carrera como un abogado y la terminó como un hereje. Después de su conversión del paganismo en 196, permaneció por solamente cinco o diez años como miembro de la Iglesia ortodoxa; tanto entonces como después de su caída en la herejía montanista produjo una serie de obras teológicas notables por su vigoroso razonamiento y desenfadado uso de la retórica legalista en contra de sus oponentes, y una aceptación intransigente de conclusiones paradójicas. La paradoja que quiero discutir proviene de una obra intitulada *De carne Christi* que escribió en el año 208, *libris*, como la *Patrología (Vit. Tert.)* lo dice de manera elegante, *iam Montanismam redolentibus* (“en un tiempo en que sus escritos ya apestaban a montanismo”) pero la obra no es en sí misma, creo, herética. Ataca a Marción, quien creía que Cristo no había, de hecho, nacido de la carne, sino que era un “fantasma” de forma humana. La negativa de Marción a creer en una genuina encarnación, argumenta Tertuliano, sólo podía provenir de una creencia de que ello sería imposible o indigno, una degradación vergonzosa de la naturaleza divina. En contra de la concepción de que sería imposible, produce el principio barredor y general *nihil impossibile Deo nisi quod non vult* (“nada es imposible para Dios, excepto lo que no desea hacer”). En particular, Marción había argumentado que la idea de la encarnación de Dios implicaba una contradicción, porque ser nacido como un ser humano implicaría un cambio en la naturaleza divina;^[25] pero un cambio implica dejar de tener algunos atributos y adquirir otros, y los atributos de Dios son eternos, por lo tanto, no pueden cambiar; de esta manera, no pudo haber nacido como un ser humano. Por su lado, Tertuliano dice que esto es argumentar falsamente a partir de la naturaleza de los objetos temporales hacia la naturaleza de lo eterno e infinito. Es por supuesto verdadero para los objetos temporales que si cambian pierden algunos atributos y adquieren otros; pero suponer que lo mismo es verdadero para Dios es simplemente dejar de lado las necesarias diferencias entre Dios y los objetos temporales (*de c. C. iii*). (Diré en la sección 6 de este artículo algo, quizá no del todo convincente, acerca de este argumento). Finalmente, en contra de la concepción de que, incluso si fuera posible, Dios no desearía encarnarse porque sería indigno de él, Tertuliano, resumiendo sus objeciones a Marción en un pasaje de gran intensidad, lo

acusa de derribar la base entera de la fe cristiana: su argumento también destruiría la crucifixión y la resurrección.

Marción elimina estos también —dice (*ibid.*, v)—, o más bien éstos: pues ¿qué es más indigno de Dios, más vergonzoso, nacer o morir? [...] Respóndeme esto, tú carnicero de la verdad. ¿No fue Dios realmente crucificado? Y como él fue realmente crucificado, ¿no murió realmente? Y como él murió realmente, ¿no se levantó de los muertos? [...] ¿Es falsa nuestra fe entera? [...] No deseches lo que es la única esperanza del mundo entero. ¿Por qué destruyes una indignidad que es necesaria para nuestra fe? Lo que es indigno de Dios, será suficiente para mí [...] el Hijo de Dios nació; dado que es vergonzoso, yo no estoy avergonzado; y el Hijo de Dios murió; justamente porque es absurdo, ha de ser creído; y fue sepultado y se levantó nuevamente; es cierto, porque es imposible. (*Non pudet, quia pudendum est... prosus credibile est, quia ineptum est... certum est, quia impossibile*).

Ésta es la paradoja de Tertuliano.

3

Las personas que se expresan en paradojas se hallan en una posición fuerte, y cuanto más escandalosa sea la paradoja, en general, su posición será más fuerte, pues un objetor que insiste en señalar lo absurdo de lo que ha sido dicho, se halla incómodamente consciente de que está haciendo el papel de tonto, pues todo lo que hace es señalar que la paradoja es paradójica, y esto ya era perfectamente obvio: es como un hombre que no entiende una broma, una observación irónica o una comparación imaginativa, e insiste en tomarla literalmente. Pero las observaciones irónicas y las comparaciones imaginativas pueden tener su razón de ser, y lo mismo las paradojas; así que no basta con que el objetor deseche la paradoja con la esperanza de que su evidente absurdidad la haga indigna de discusión, pues esto es nuevamente sugerir que la persona que la ha expresado ha pasado por alto su absurdidad cuando, por el contrario, él sabe que era absurda, y esa era una de las razones por las que la expresó. Debido a que las personas, en general, no expresan absurdos a menos que tengan una razón para hacerlo, se siente que el paradoxógrafo debe haber dicho algo importante. No solamente evita la respuesta de los críticos, sino que les hace sentir que en algún modo misterioso se halla en una mejor posición que aquellos; es más bien como un hombre normalmente bien vestido que aparece en una función con una corbata negra y frac: los otros presentes no se lo pueden mencionar, no pueden dejar de verlo, y se sienten incómodos frente a su propio rechazo. O, nuevamente, él es algo así como un hombre que cierra firmemente una puerta en la cara de uno: no sólo evitando que uno entre, sino haciéndole sentir que no tiene derecho a hacerlo.

Hasta aquí, el paradoxógrafo tiene todo de su lado, pero no enteramente. Pues, como el hombre con la corbata negra, para provocar este efecto tiene que estar por lo regular bien vestido, y el hombre que cierra la puerta tiene que ser alguien que uno respeta, así el paradoxógrafo debe tener alguna otra pretensión o algún otro reclamo de la atención de su público: pues, en general, una paradoja, no importa cuán sugerente sea en sí misma,

no representa ganancias sólidas —toma un poco del crédito de ayer o hipoteca un poco el de mañana—. Este derecho a la atención de uno puede ser poseído de varios modos: positivamente, si el que se expresa es un persona genuina, buena e impresionante, cuya vida demanda amor y respeto, o por otras expresiones tuyas que hayan sido originales y profundas, y negativamente, por otras exigencias a nuestra atención que sean contradictorias, o aparentemente contradictorias, que sean confusas o inútiles, o hechas por personas cuyo modo de vida parece trivial, malo o desastroso. Si esto es así, podríamos esperar encontrar que las creencias de una religión, por ejemplo, son presentadas con una paradójicalidad particularmente desafiante en dos tipos de situaciones: primero, cuando sus creyentes se hallan intensamente vinculados por una fe nueva y convincente, y peleando por su supervivencia en una sociedad hostil pero decadente cuyas creencias rechazan completamente, y segundo, cuando no importan las divisiones y descrédito que hayan caído sobre la creencia misma; aquéllos que la rechazan, estando sus propias esperanzas pereciendo, al parecer tienen poco que ofrecer en su lugar, excepto *congoja*, tiranía o una inminente aniquilación termonuclear.

Sin embargo, lo dicho hasta aquí sugiere solamente por qué las personas, y en particular los creyentes, deberían tender algunas veces a expresarse en paradojas; dice poco acerca de por qué nadie debería, en absoluto, optar por hablar en paradojas, o sugiere, a lo sumo, que lo hacen como un modo de llamar la atención para hacer que la gente escuche o considere algo más. Frecuentemente, no es mucho más: decir, por ejemplo, que el Sacro Imperio Romano no era sacro ni imperio ni romano es, o debería ser, una manera enérgica de preparar un nuevo análisis histórico. Pero hay otras paradojas que parecen más importantes y significativas; donde captar la paradoja parece una parte esencial de entender lo que se está diciendo. Aquí tenemos la sensación de que una paradoja, concediendo que tiene que entenderse en un trasfondo de otras creencias o un modo de vida, en sí misma nos dice algo: que es en cierto modo la esencia de lo que ha de creerse. Esto es particularmente así en el caso de las creencias religiosas, donde la sensación ha sido expresada de muchos modos: quizá diciendo que hay una infinidad de cosas que se hallan más allá de nuestra comprensión; o que nuestra razón no puede abarcar las verdades más profundas; o que lo que decimos sólo puede ser una explicación insatisfactoria (o, quizá, analógica) de lo que creemos por fe. Trataré de mostrar cómo tal punto de tensión, de falla del lenguaje, debe ocurrir en la creencia religiosa, y pienso, por lo tanto, que debiéramos tomarnos en serio la paradoja de Tertuliano; no sólo como una expresión retórica de sus objeciones a una doctrina particular, sino como una impactante formulación de algo que habré de sugerir es esencial a la fe cristiana.

4

En años recientes se ha discutido mucho sobre el lenguaje religioso y sus relaciones con otros tipos de lenguaje; buena parte de esta discusión se ha concentrado en las *proposiciones* religiosas, y de ésta, una gran porción en la de “Dios existe”. Pienso que

es tiempo de considerar si tal concentración no ha sido demasiado estrecha: pues en cada aspecto ha tenido resultados indeseables. Primero, hay una falta de claridad en la idea de un lenguaje. No me refiero, desde luego, a un lenguaje nacional o un lenguaje de diccionario, tal como el francés o el esperanto, sino a un lenguaje o tipo de discurso lógicamente distinguible. Segundo, la concentración de las proposiciones religiosas, en tanto que distintas de otros tipos de expresión religiosa, ha producido una cadena de efectos perturbadores: ha dado demasiada importancia a la diferencia entre la proposición religiosa aparentemente infalsificable y la proposición falsificable de las ciencias, la cual es, desde luego, importante y aparecerá posteriormente en lo que tengo que decir, pero que tomada por sí misma conduce a un *impasse* que parece una reduplicación en términos lingüísticos de la estéril disputa decimonónica entre ciencia y religión, y los esfuerzos por salir de este *impasse* han implicado, en algunos casos, intentos de reducir las proposiciones religiosas a proposiciones de otro tipo, por ejemplo, de experiencia mística, y en otros intentos de reducirlas a otras cosas que no son proposiciones en absoluto, tales como órdenes o exhortaciones a un modo de vida religioso, todos los cuales o bien implican una evidente circularidad u omiten completamente el carácter peculiarmente religioso de las proposiciones. Tercero, se ha puesto especial atención a la lógica de la proposición particular “Dios existe”. Esto muestra una especie de valentía sin esperanza. Muestra valentía porque esta proposición parece ser la clave de todo el sistema: descubrir lo que está implicado en creer esto debería ser descubrir la naturaleza entera del lenguaje religioso y la esencia de la creencia religiosa. Pero es precisamente la importancia peculiar de esta proposición lo que deja sin esperanza una investigación que empieza con ella. Su peculiaridad la hace extremadamente atípica dentro de las proposiciones religiosas; una peculiaridad enfatizada por Collingwood, por ejemplo, cuando dijo que no era una proposición religiosa en absoluto, sino más bien la presuposición de cualquier proposición religiosa. Podríamos decir que la proposición de la existencia de Dios tiene, desde luego, un gran poder lógico, pero que es el poder no tanto de un perno como de una palanca: si supiéramos, desde fuera del sistema religioso, cómo operar con ella, podríamos mover el cielo y la tierra, pero no podemos hacerlo así, porque no sabemos ni dónde podemos fijar un fulcro ni dónde podemos insertar el otro extremo de la palanca. Así que, más que intentar tal aproximación directa, debemos reconocer la dificultad y darle la vuelta.

Espero no dar un rodeo muy largo, pero quizás algo se pueda decir. Primero, pienso que debemos siempre tener en mente el hecho de que el lenguaje religioso no se usa simplemente para hacer proposiciones, sino que hay muchos otros tipos de expresión religiosa: órdenes, por ejemplo (“No tomarás el nombre de Dios en vano”), y, de modo muy importante, oraciones y expresiones de confianza (“Aunque él me destruya, sin embargo lo esperaré”); también promesas y reprimendas, y muchas otras. Más aún, ninguna de estas expresiones, incluyendo las proposiciones, se hace *in vacuo*: algunas veces se utilizan como parte de una ceremonia u observancia religiosa; otras como parte del proceso en el que una persona religiosa debe decidir qué hacer en una situación práctica, y generalmente como parte de las actividades de la vida. Esto, como un punto

general, es constantemente enfatizado por Wittgenstein, y al considerar el lenguaje religioso es, pienso, particularmente desastroso ignorarlo.

Pero, ¿qué es el lenguaje religioso? ¿Hay algo que se pueda llamar lenguaje religioso? ¿Con qué se está contrastando? Por cierto, algo con lo cual debo ser cauteloso al comparar es esa sustancia nebulosa y omnipresente llamada “lenguaje ordinario”. Por algo, el lenguaje ordinario debería ser el lenguaje utilizado por la mayoría de nosotros en nuestra vida diaria, y la cuestión de cuán religioso es ello, es la pregunta de cuán religiosos o declaradamente religiosos somos la mayoría de nosotros; y si algunos de nosotros nos abstenemos siempre, y la mayoría de nosotros la mayor parte del tiempo, de hablar acerca de Dios al tratar nuestros asuntos, eso parece decir de nosotros al menos tanto como del hablar acerca de Dios. Esto plantea la cuestión de prescindir de hablar acerca de Dios, de lo que está implicado en hacerlo sin tocar el tema. Sobre este punto hablaré más adelante.

Así que uno podría preguntarse cuáles son, en general, las marcas distintivas de un lenguaje o de un tipo de discurso, y al intentar un bosquejo de respuesta, uno podría pensar de inmediato al menos en cinco marcas distintivas. Un lenguaje puede distinguirse de otro por los tipos de relación lógica que tienen lugar en él: por su tema, por su uso de términos técnicos, por sus propósitos, o, más generalmente, por las actividades con las cuales se halla asociado. Pero sería, desde luego, una ilusión suponer que estas cinco marcas, incluso si fueran satisfactoriamente distintas unas de las otras, pudieran competir por la posición de marca distintiva única entre un lenguaje y otro. Podemos decir, más bien, que de la interrelación de características como éstas uno puede, en casos particulares, justificadamente pretender distinguir un tipo de discurso de otro. A cuál de estas características se podría dar una atención particular depende en parte del nivel en el que se estén trazando las distinciones. Si nos concentramos en las distinciones entre las ciencias, a un bajo nivel de generalidad, tendemos a aferrarnos a distinciones de tema, pues todos aprendemos en la escuela que la micología es el estudio de los hongos, y la geología el estudio de las rocas, etcétera. Pero está claro que al hacer esto presuponemos ya una distinción entre el discurso científico y otro, y entre un tipo de ciencia y otra: pues no toda mención acerca de las plantas, por ejemplo, y ni siquiera todo discurso científico acerca de ellas es botánico. Ni la distinción de tema se aplicaría a cualquiera, sino a las más ingenuas distinciones entre ciencias subsidiarias; la distinción, por ejemplo, entre física y química física, se halla más bien en el alcance y la terminología de las leyes formuladas y empleadas.

Pero no es mi propósito, aunque fuera posible, intentar la elevada tarea aristotélica de caracterizar las diferencias entre cuerpos organizados de conocimiento, pues mientras que puede caracterizarse el lenguaje utilizado por algún tipo de científico en su trabajo profesional, o una actividad científica profesional para distinguirla de otra, tal como la del historiador, esto está fuera de nuestro propósito, el cual es caracterizar algunos usos no profesionales del lenguaje en tanto que distintos de otros. Parece claro, sin embargo, que cuando nosotros, como legos, especulamos sobre la distancia de una estrella, por ejemplo, estamos usando el lenguaje de modo diferente de cuando observamos cuán

bellamente brilla ahora, y que si decimos que el primero es un uso científico, parte de que lo que queremos decir es que estamos haciendo una pregunta a la cual el científico profesional se halla en la mejor posición de dar una respuesta —es el tipo de pregunta que él se hace—. Así que podemos en este punto reintroducir la idea de un uso profesional del lenguaje y decir al menos que algunas de nuestras expresiones plantean o implican preguntas que son propiamente contestadas por técnicas y métodos de investigación profesionalmente empleados por algunos tipos de especialistas, y otras no lo son.

No es de ningún modo simple aplicar esta distinción a nuestra investigación del lenguaje religioso. En el caso de la creencia religiosa hay, desde luego, la noción de que una persona es una autoridad, pero esto es algo muy diferente de una autoridad científica. Pues, primero, la autoridad religiosa, si es que hay una, es al menos no simplemente alguien que tiene un buen entrenamiento en los métodos de responder ciertos tipos de preguntas, sino alguien que *tiene* la autoridad para *establecer* lo que ha de creerse o hacerse. Segundo, la cuestión de si hay una autoridad religiosa incluso en este sentido y, en tal caso, quién es esa autoridad, ha dado ocasión a disputas violentas, y muchas personas han sido muertas en los intentos por dirimirla, pero el tema que nos ocupa son las preguntas admitidas como religiosas, así que una referencia a la autoridad no puede venir de la caracterización de una pregunta religiosa. Tercero, incluso si fuéramos a decir que un uso especialista o profesional del lenguaje religioso se pudiera encontrar en su uso teológico (y acerca de la teología tengo algo que decir después), está claro que la relación del lenguaje religioso con el teólogo es diferente de la relación entre el lenguaje de la ciencia y el científico; alguien que habla científicamente es al menos un científico amateur, pero quien habla religiosamente no es necesariamente un teólogo, ni siquiera un aficionado.

¿Cómo podemos, entonces, intentar una caracterización de las expresiones religiosas, o al menos cristianas, de personas que no sean teólogos profesionales en su vida cotidiana? ¿Podríamos decir, por ejemplo, para adoptar uno de los criterios sugeridos, que cierto lenguaje ha de ser caracterizado como religioso, o más específicamente como cristiano, por referencia a ciertas prácticas u observancias en el curso de las cuales se usa? Está claro que tal referencia no nos daría demasiado, pues en primer lugar las ceremonias tendrían que ser en sí mismas caracterizadas como ceremonias religiosas y, si podemos hacer eso, deberíamos ya tener una idea más clara de lo que es el lenguaje religioso. Segundo, muchas expresiones religiosas se hacen fuera de tales ceremonias, a menos que uno hable religiosamente solamente los domingos. Las ceremonias podrían al final tener que ser mencionadas en una plena caracterización de una vida religiosa, pero lo que estamos buscando primero debe encontrarse en otra parte. Hemos visto que, al intentar la caracterización de otros tipos de lenguaje, la distinción del tema en el cual dicho lenguaje se utiliza, no nos lleva muy lejos, pero en el caso del lenguaje religioso quizá deberíamos, después de todo, regresar a ello, pues éste es, podríamos decir, propiamente el lenguaje acerca de Dios; y por “propiamente” quiero decir no sólo que todo el lenguaje religioso es lenguaje acerca de Dios, sino —y esto me parece un punto

importante— que todo el lenguaje acerca de Dios es lenguaje religioso.

Pero decir que el lenguaje religioso es lenguaje acerca de Dios inmediatamente plantea tres dificultades relacionadas, pues en primer lugar el término “acerca de” es equívoco. En el sentido lingüístico más normal de “acerca de”, son las proposiciones las que se refieren a cosas o personas; pero, como ya hemos visto, no todas las expresiones religiosas son proposiciones; una oración, por ejemplo, no es acerca de Dios, sino que está dirigida a él. Si hemos de decir, entonces, que el lenguaje religioso es lenguaje acerca de Dios, tenemos que tomar “acerca de” en un sentido extremadamente amplio. Supongo que no sería disputado por los cristianos que toda expresión religiosa, en algún sentido, regresa a Dios; quizá en el sentido de que si el propósito de la expresión ha de ser explicada, Dios tiene que ser mencionado en última instancia. Más o menos en este sentido debe entenderse el término “acerca de”. Pienso que tenemos que decir, más aún, que la mera ocurrencia de la palabra “Dios” en una expresión no significa que ésta sea realmente acerca de Dios, y por ende religiosa, ya que los más devotos pueden usar la palabra “Dios” en frases ociosas y no pretender realmente hablar acerca de Dios. Una expresión que incluye la palabra “Dios” debe pretender seriamente ser acerca de él para que verdaderamente sea acerca de él.

A la inversa —ésta es la segunda dificultad— no es el caso que la palabra “Dios” *tenga* que figurar en una expresión para que ésta sea religiosa. Podríamos formular esto diciendo que la distinción de tema no se puede reducir a otra distinción que mencioné, la de terminología técnica, pues hay muchas expresiones que son religiosas, incluso en el sentido de “cristiano”, pero que no implican el uso de la *palabra* “Dios”; y, más ampliamente, hay un lenguaje religioso que no es el lenguaje de la cristiandad. Decir que este otro lenguaje religioso es lenguaje acerca de Dios, donde “Dios” se entiende en un sentido cristiano, es por lo menos prejuzgar un asunto teológico particular, concerniente a la referencia y verdad de las creencias religiosas que no son la cristiana; pero el hecho de que haya un resultado aquí, muestra que debe haber alguna caracterización de una religión y, con ello, del lenguaje religioso que es independiente de las creencias de la cristiandad. De este modo, tenemos que decir que no caracterizamos el lenguaje religioso en general, sino específicamente el del cristianismo, y eso es suficiente para los propósitos de esta discusión.

Pero ¿es al menos esto? Pues —y esta es nuestra tercera dificultad— decir que el lenguaje cristiano es lenguaje acerca de Dios evidentemente presupone la verdad de la cristiandad en un sentido mucho más radical, ya que presupone la existencia de Dios. Por lo tanto, parece como si tuviéramos que decir que, si Dios existe, el lenguaje de la cristiandad es un lenguaje acerca de Dios, y esto parece inútil como una caracterización de tal lenguaje, pues si empezamos a partir de la proposición de la existencia de Dios, la caracterización sería vacua a menos que ya sepamos que esa proposición es verdadera, y si sabemos que esa proposición es verdadera, la caracterización parecería superflua. No obstante, si empezamos a partir de la evidente existencia de un lenguaje cristiano, en el sentido del lenguaje utilizado por los cristianos, podríamos estar tentados a llegar a la proposición de la existencia de Dios, y así implicarnos en una especie de prueba

ontológica que podría muy bien ser considerada sospechosa. Todo esto ilustra la peculiar relación del lenguaje religioso con la proposición de la existencia de Dios, que ya he mencionado. Si estuviéramos buscando una caracterización independiente del lenguaje cristiano, estas dificultades serían críticas, pero mi intención no es hacer esto, sino dejar por un lado la cuestión de la existencia de Dios, y tratar de mostrar algo acerca del lenguaje cristiano tal y como lo usan los cristianos. Así que quizá esta aproximación, más bien paradójica, no resultará ser enteramente inútil. Por lo tanto, continuaré hablando del lenguaje cristiano sin tomar en consideración la incredulidad, lo que se logra, evidentemente, por nuestro propio esfuerzo. Esto es, sugerir que todo el lenguaje cristiano es lenguaje acerca de Dios, y ya antes había dicho que todo lenguaje acerca de Dios es lenguaje religioso, y esto debe mantenerse, si es que se mantiene, en su forma original: pues decir que todo lenguaje acerca de Dios es lenguaje *cristiano* es prejuzgar para el efecto opuesto el asunto teológico, que ya mencioné antes, acerca del estatus de otras religiones.

Pero aquí tal vez tengamos un punto importante acerca del lenguaje religioso: pues vimos antes que, mientras que el lenguaje de la botánica es lenguaje acerca de las plantas, no todo el lenguaje acerca de las plantas es botánico, pues los poetas, pintores, vagabundos, etc., pueden hablar también de plantas, pero no botánicamente. Quiero sugerir que todo el lenguaje acerca de Dios es lenguaje religioso; uno no puede hablar no religiosamente acerca de Dios. No se deduce de esto que los ateos se hallan necesariamente hablando religiosamente: ellos están negando la proposición “Dios existe”, y hacer esto queda fuera de las presuposiciones de la presente discusión. Se deduce, sin embargo, de la tesis presente, que la blasfemia es una especie de lenguaje religioso y debe ser tal —pues ¿de que otra manera podría ser tan ofensivo?—. La blasfemia es el mal uso del lenguaje religioso: es decir cosas acerca de Dios, o preguntar cosas de él, pero las cosas equivocadas. No obstante, parece haber una especie de paradoja aquí: pues el blasfemo dice, por ejemplo, que Dios es malvado, y al hablar así del Dios cristiano, ofende, pero el Dios cristiano es bueno; así que, ¿no debería el blasfemo estar hablando de otro Dios? Pero si él hace esto, o bien no ofende, porque no es el Dios cristiano del que estaba hablando ofensivamente, u ofende solamente al sugerir que hay otro Dios, una línea de argumento que podría conducir a la intrigante conclusión de que la única forma de herejía es el politeísmo. Quizá aquí deberíamos decir algo como lo que decimos acerca de los desacuerdos concernientes a los personajes en el pasado histórico, que debe haber algunas creencias, y en el caso de la religión algunas prácticas, en común entre el blasfemo y el ortodoxo para apoyar la idea de que ambos están hablando del mismo Dios: cuando Housman se refería a “Cualquier tonto o pillo que hizo el mundo”, la descripción “Él, quien hizo el mundo” proveyó el lugar del cual la ofensa había de ser tomada.

Si decimos, entonces, que todo lenguaje acerca de Dios es lenguaje religioso, hemos dicho algo acerca del lenguaje religioso y su tema, que lo distingue de, digamos, el lenguaje botánico y su tema. En seguida, debemos considerar un tipo de expresión que, de manera muy importante, ocurre dentro del cuerpo del lenguaje religioso. He

enfaticado el hecho de que las expresiones religiosas pueden ser de muy variados tipos: proposiciones, órdenes, oraciones, etc., pero es importante distinguir también entre distintos tipos de proposiciones. Algunas pueden ser proposiciones directamente acerca de la naturaleza de Dios: “Dios es tres personas”; otros, de eventos históricos: “Dios envió a los judíos al exilio”, y otros más, de la naturaleza humana: “Dios ha dado al hombre libre albedrío”, etc. Es decir, hay muchas proposiciones religiosas que no son simplemente proposiciones religiosas; aunque hablen acerca de Dios, se refieren también a otra cosa, a algo que afecta los asuntos de los hombres.

5

Debemos ahora mirar más de cerca el modo en que algunas proposiciones religiosas, para restringir nuestra investigación solamente a las proposiciones, no hablan solamente de Dios, sino también de los asuntos humanos, pues al hacerlo así podemos alcanzar más claridad acerca del alcance del lenguaje religioso, y de su relación con otros lenguajes y con la teología: y regresaremos, finalmente, a la paradoja de Tertuliano. Debido a que las proposiciones religiosas son tan variadas, deberíamos analizar muchas de ellas, pero aquí solamente mencionaré una. Habla de algunos asuntos muy trillados, pero sólo lo tomo como ejemplo para ilustrar un punto más general.

Si un pueblo tiene una ocasional falla en sus cosechas, con la subsecuente hambruna, alguien guiado por el Antiguo Testamento podría decir: “Dios hace que las cosechas se pierdan para castigar a las personas por su maldad”. Tal proposición es, ciertamente, una especie de proposición religiosa, pero también es una proposición acerca de algunos eventos en la vida humana, y parece, de hecho, proveer una explicación de los mismos, y lo hace de la manera más clara, lógicamente similar a una explicación no religiosa de los mismos eventos, porque conecta entre sí dos conjuntos de eventos humanos —la maldad de la gente y la pérdida de sus cosechas— con Dios, como quien dice, como intermediario. En sí misma, la proposición parece ser también en un sentido fraudulenta, pues cuando llegan los agricultores, mejora la irrigación, las cosechas nunca se pierden y la gente se regodea en su maldad en medio de la abundancia; el hombre que dijo que las cosechas se perdían debido a la maldad de las personas cae notoriamente en descrédito. Se dejarán de atribuir a Dios el éxito o fracaso de las cosechas: una especie de proposición religiosa que ya no se hará. Éste es el familiar fenómeno de la eliminación del lenguaje religioso de un contexto, y aquí ha sido eliminada no sólo porque se deja de hablar ya de cierto modo —de la misma manera que alguien podría dejar de escribir algún tipo de poesía— sino porque la proposición religiosa, en este particular ejemplo, era una especie de explicación, y fue superada por otra rival y mejor. Mencionamos antes una distinción de lenguajes por diferencias técnicas y de especialidades, y en el que, cuando el lenguaje del especialista puede chocar con el lenguaje religioso, la tendencia es descartar este último, debido a que las técnicas del especialista dan explicaciones que son recomendaciones para la acción efectiva, y a que, donde el lenguaje religioso pretendía

hacer eso, fracasó: pues o bien daba una “explicación” que no lo era en absoluto, y no daba ninguna recomendación, o daba una explicación, como en el caso de las cosechas, pero muy mala, pues cualquiera que crea que la mejor manera de evitar los desastres naturales es vivir una mejor vida, al parecer está en un error.

Sería un burdo error, desde luego, suponer que estas antiguas consideraciones deberían, en algún sentido, “refutar la religión”. Lo que muestran es que si el lenguaje religioso es usado para dar cierto tipo de explicaciones, choca con una explicación más efectiva y tiende a ser eliminado. Tal eliminación tiene también su efecto en la teología del usuario del lenguaje, pues la explicación religiosa, como hemos visto, era una proposición tanto acerca de ciertos eventos como acerca de Dios, y si estas proposiciones son vistas como inaplicables a los eventos, son vistas como inaplicables también a Dios. Por ende llegará a verse, quizá, que ciertas cosas no pueden decirse de Dios: por ejemplo, que produce desastres particulares como un castigo para los hombres. Esto, a su vez, conduce a una nueva especulación acerca de la naturaleza de Dios y su relación con el mundo, de modo que un cambio en los posibles usos del lenguaje religioso está conectado con un cambio en las concepciones acerca de la naturaleza de Dios. Esto funciona también del otro modo, pues sería falso representar la situación de la que hablamos como una constante retirada del lenguaje religioso, con un consecuente recorte de la doctrina teológica. Indudablemente, esto ha sucedido, pero puede haber también nuevos pensamientos acerca de Dios y nuevas concepciones morales que resultan de éstos; cambios en lo que puede decirse en el lenguaje religioso; como con la llegada del cristianismo se dijo menos acerca de la ira de Dios y más acerca de su misericordia. Con ello viene también un cambio en la creencia religiosa, un cambio en lo que es considerado propio pedirle en oración. En cada etapa se puede hacer un intento de ordenar las implicaciones de lo que los hombres piensan que pueden y deberían decir en el lenguaje religioso, y tal orden asume la forma de una serie de proposiciones acerca de la naturaleza de Dios: y esto es teología sistemática.

Pero, aunque los cambios en el alcance del lenguaje religioso no han de ser descritos enteramente como una retirada, la retirada, como todos sabemos, tiene sus peligros para la religión. La supuesta explicación religiosa que mencionamos fue, en su forma general, una proposición acerca de Dios y el mundo, y si se abandonaran todas las proposiciones sobre Dios y el mundo, ¿qué quedaría? Tales proposiciones no tendrían que ser, como esa lo era, explicativas de lo que ocurre en el mundo, y, desde luego, no podrían serlo; pero hay conexiones que no son explicaciones, y algunas deben seguramente serlo. Wittgenstein dijo (*Tractatus*, 6.432): “Cómo sean las cosas en el mundo es un asunto completamente indiferente para lo más alto. Dios no se revela a sí mismo *en* el mundo”. Pero si todo lo dicho acerca de Dios fuera solamente acerca de Dios, y todo lo dicho acerca del mundo fuese sólo acerca del mundo, ¿cómo podría ser que Dios fuese el Dios *del* creyente cristiano, quien se halla activo en el mundo de los hombres? Las concepciones acerca de la naturaleza de Dios se retirarían más y más del mundo de los hombres: su existencia se volvería como la de los dioses de Epicuro, “muy remotos y separados de nuestros asuntos” (Lucrecio, de *Rerum Natura*, II, 648). Y si ello ocurriese,

no tendría mucha importancia si se hallaran ahí o no.

6

Aquí es donde regresamos a Tertuliano. La paradoja de Tertuliano es relevante para nuestra pregunta porque es una paradoja y porque es acerca de la encarnación. Pues la encarnación parece ser el punto de la fe cristiana donde debe haber esencialmente una intersección entre el lenguaje religioso y el no religioso; tiene que decirse no solamente que una cierta persona fue crucificada, sino que esa persona era el Hijo de Dios. Esto tiene que decirse, como lo vio claramente Tertuliano, si es que ha de haber una fe cristiana; y como él lo vio de manera igualmente clara, es una paradoja. La paradoja aparece porque, aunque nosotros debemos tener alguna proposición que diga algo acerca tanto de Dios como del mundo, cuando lo tenemos encontramos que tenemos algo que no podemos decir propiamente, pues cuando se habla de Dios en un lenguaje puramente religioso, se dice que es eterno y perfecto; esto es, que nosotros no podemos hablar de él apropiadamente en los términos que empleamos para los objetos temporales imperfectos y las personas de esta vida; o si lo hacemos, es bien conocido por la analogía de la cual hablan los teólogos y, por lo tanto, de forma imperfecta, pues no hay lenguaje para la eternidad y la perfección de Dios fuera de la proposición que las afirma: se puede decir *que* Dios es eterno y perfecto, pero no *cómo* es él, pues la eternidad de Dios y su perfección deben estar más allá del alcance de nuestro entendimiento. Así que cuando llegamos a una proposición acerca de Dios y de los sucesos temporales debe ser insatisfactoria, pues si no lo fuera deberíamos haber descrito adecuadamente la relación entre los sucesos temporales y Dios en términos apropiados sólo para los sucesos temporales: y esto significaría o bien que hemos descrito solamente los sucesos temporales y dejado a Dios de lado, o bien que hemos descrito a Dios como un ser temporal, lo cual no es.

La dificultad parece surgir no de la eternidad de Dios en sí misma, sino de la conjunción de ésta con su perfección como un ser personal, pues algunos han mantenido, por ejemplo, que los números son objetos eternos, aunque las proposiciones matemáticas acerca de las cosas en el mundo se puedan hacer de manera satisfactoria. Una diferencia entre este caso y el de Dios es, dejando de lado la cuestión de la aplicación, que la naturaleza de los números en sí mismos se puede expresar adecuadamente en el lenguaje de las matemáticas puras, pero la naturaleza de Dios no puede ser expresada adecuadamente en ningún lenguaje humano. Pero si decimos esto, parece como si estuviéramos defendiendo ahora una tesis diferente acerca del lenguaje religioso, pues esto parece decir que cualquier proposición acerca de Dios, ya sea que digamos que hay una relación entre Dios y el mundo o no, sería insatisfactoria, justamente porque está hecha en las palabras del lenguaje humano. Sin embargo, la tesis era que es el hecho de que debe haber una relación entre Dios y el mundo lo que hizo insatisfactorio el lenguaje religioso, pero no es realmente una tesis diferente, pues es justamente el hecho de que

hay en algún punto tal relación, y una proposición o conjunto de proposiciones que traten de expresarla, lo que hace que el lenguaje religioso sea insatisfactorio también aquí. La cuestión de la aplicabilidad de las matemáticas al mundo no afecta la cuestión de la expresión de la naturaleza de los números por las matemáticas puras, pero la cuestión de la relación de Dios con el mundo afecta la cuestión de la expresión de la naturaleza de Dios en el lenguaje religioso. El efecto real es que se dice que Dios es un ser personal perfecto, porque, por ejemplo, se le dirigen oraciones, y porque él tiene un Hijo que nació en el mundo. La proposición de estas relaciones sería en sí misma insatisfactoria, e implicaría otras que también lo son, porque los conceptos que requiere —de paternidad, por ejemplo, de amor y de poder— son adquiridos en un contexto humano; es un lenguaje que surge y se utiliza para las relaciones de los humanos con los humanos. Decir que, mientras que esto es así, el lenguaje religioso requiere meramente una extrapolación del contexto humano[26] no es resolver el problema sino volver a plantearlo, pues la extrapolación requerida es una extrapolación al infinito, e incluso al tratar de darle un sentido nos encontramos con lo incomprendible. Es esta incomprendibilidad la que Tertuliano ha destacado en su paradoja.

De hecho, es una doble paradoja: “porque es vergonzoso”, dice Tertuliano, “no estoy avergonzado [...] es cierto, porque es imposible”; esto es, hay algo en eso que es moral e intelectualmente escandaloso. Las dos paradojas quizá puedan ser vistas considerando la encarnación desde dos diferentes direcciones. Que Dios, un ser perfecto, quisiera nacer y ser crucificado, es moralmente sorprendente; que este hombre sobre una cruz debiera ser realmente Dios es intelectualmente sorprendente. De éstas, la paradoja moral es quizá la más rápidamente comprensible para el incrédulo, pues ésta, al menos, tiene como modelo las ideas de humildad y sacrificio, y la búsqueda del mayor valor donde el mundo no lo está buscando. Así que el incrédulo, quizá de manera impertinente, puede sentir que ve un punto en la paradoja moral, que ha puesto al revés las normas de lo que ha de ser admirado y amado. El sentimiento es más fácil, tal vez peligrosamente, porque tenemos una tradición cristiana; para el romano instruido, por ejemplo, debe haber sido profundamente escandaloso. En este caso, también, podemos entender hasta cierto punto lo que es el centro de la paradoja, el hecho de que Tertuliano no dijera que era absurdo y que él lo creyera, sino que él lo creía *porque* era absurdo. Era precisamente lo escandaloso de la crucifixión lo que apuntaba al nuevo camino que uno tenía que tratar de seguir en la propia vida,[27] y ¿cómo pueden cualesquiera de éstas cosas ser aplicadas a la segunda parte de la paradoja, al hecho de la encarnación? ¿Cómo puede ser cierto, porque es imposible? ¿Cómo puede llegar a entenderse, cómo se le puede dar cualquier sentido a la proposición de que este hombre que fue crucificado era Dios?

Aquí encontramos claramente una dificultad que poco a poco se ha dejado sentir a lo largo de esta disquisición, pues el examen del significado de las proposiciones acerca de la encarnación es, o ha sido, una tarea para teólogos, y dado que no soy un teólogo, no me siento competente para abordarlo. No obstante, al abordar el lenguaje religioso al parecer he alcanzado un punto en el que es necesario convertirse en teólogo. Pienso que podemos ver la razón de esto si consideramos lo que se ha dicho ya acerca del lenguaje

religioso y la teología. Sugerí antes que había una relación entre lo que puede y no puede decirse en un lenguaje religioso y la teología sistemática; que una contracción o extensión en el uso del lenguaje religioso conducía a cambios en la teología, y que la sistematización y explicación de las implicaciones de lo que puede decirse acerca de Dios es una tarea de la teología. Esto, sin embargo, parece tener dos consecuencias, pues la teología examina los resultados lógicos de hablar de este u otro modo acerca de Dios y cambia con ellos: si uno no puede decir “Dios envió la sequía para castigar a la gente”, debemos decir que Dios no interviene en las operaciones de la ley natural; si decimos esto, ¿hemos de decir que el poder de Dios está limitado o que él mismo no quiere intervenir?, y así sucesivamente, todos los problemas tradicionales de la teología. Asimismo, si decimos que Dios se encarnó, ¿hemos de decir que cambió? (el problema de Marción). Pero si el planteamiento de esas cuestiones es una tarea de la teología, entonces la teología parece incluir el análisis lógico del lenguaje religioso; pues seguramente el análisis lógico del lenguaje religioso es justamente esto, preguntar cómo, y con qué implicaciones, se profieren expresiones en contextos religiosos. Así que el filósofo que considere que su tarea consiste en el análisis lógico del lenguaje, y que se proponga examinar el lenguaje religioso, sospecho que terminará, como yo lo he hecho, haciendo teología. Esto, que es la primera consecuencia, no me parece demasiado afortunada para el supuestamente independiente analista del lenguaje religioso, pues sospecho que como teólogo tendrá una actuación más bien pobre, como algunos claramente la han tenido, suponiendo que están planteando, por primera vez, dificultades lógicas en el lenguaje cristiano que de hecho en una u otra guisa han sido la preocupación de los teólogos por siglos. Si él no es un creyente en absoluto, su caso será aún peor, pues las expresiones no están simplemente *allí* para ser despedazadas sin entendimiento del contexto en que se utilizan; pero ¿puede entender cuál es el contexto e importancia de una oración, por ejemplo, si no entiende qué es orar? Más de lo que puede escribir en el lenguaje de la estética un hombre que no puede ver más allá de una fotografía a colores de “El caballero sonriente”.

Sugiero, entonces, que la primera consecuencia que lo dicho tiene sobre la naturaleza de la teología es que no puede hacer un análisis lógico independiente del lenguaje religioso, y que la segunda consecuencia es su opuesto: que si una tarea de la teología es tal análisis, entonces está comprometida a volverse coherente, no sólo consigo misma, sino también hacia afuera. Ya hemos visto cómo el lenguaje religioso puede separarse de los asuntos humanos hasta una lejanía epicúrea y que esto no debe suceder, si es que dicho lenguaje ha de tener alguna utilidad. Así, la teología debe mostrar cómo el lenguaje religioso puede engranarse con otro lenguaje, y debe exhibir los puntos de intersección. No obstante, al final parece que no puede tener éxito en esto, pues esos puntos de intersección, como he tratado de decirlo, deben contener algo incomprensible. Al decir esto, sólo me refiero a lo que los teólogos y otras personas religiosas casi siempre han dicho, y esto muestra que, en cualquier caso, se deduce de la naturaleza de la cosa que, mientras uno debiera ser un creyente para ser un teólogo, ser un creyente no elimina la incomprensibilidad, pues si la creencia es verdadera, es una creencia en un Dios eterno

pero personal con una preocupación por el mundo, y de esto se deduce la incomprendibilidad.

Puesto que acabo de descalificarme a mí mismo como posible teólogo, no perseguiré la lógica de las proposiciones de la encarnación. Diré, sin embargo, que está claro que uno no puede tratar con las dificultades del modo sumario que, en el trabajo en discusión, Tertuliano adopta. Se recordará que Marción había dicho que si Dios se hubiera encarnado, hubiera cambiado; pero el cambio implica perder algunos atributos y ganar otros, y Dios no puede hacer esto. Tertuliano vigorosamente replicó que lo que Marción había dicho de los objetos temporales era verdadero, pero que Dios no es un objeto temporal, y que, por lo tanto, lo que Marción dijo no se aplicaba. Pero esto es responder al movimiento del oponente haciendo astillas el tablero de ajedrez. Pues la objeción de Marción, podríamos decir, es un punto acerca de la lógica de la palabra “cambio”, y sólo entendemos la palabra “cambio” en términos del perder o ganar propiedades temporales: así que ¿cómo podemos utilizarla en el caso de Dios? Entonces debe decirse algo más, pero nuevamente, si las creencias son verdaderas, nada se puede decir que realmente funcione. La paradoja de Tertuliano es también una paradoja de la teología: parece comprometida en algo que en sus propias premisas debe ser una tarea imposible.

7

Si es imposible, ¿qué ha de hacerse? Aquí se puede decir que debemos tener fe, y más aún, que la incomprendibilidad que he estado discutiendo no solamente es una característica necesaria de la creencia cristiana, sino necesaria para ella, pues es esto lo que provee un lugar para la fe. En mi libre interpretación supongo que el mismo Tertuliano sugiere esto en el núcleo de su paradoja: es cierto, dice, porque es imposible.

Debemos distinguir aquí los diversos significados de “tener fe”. Podemos tener fe en una persona, en el sentido de que confiamos en su honestidad, buenas intenciones, sabiduría, etc., quizás a pesar de una aparente perversión en sus acciones. También podemos creer por fe en una proposición sobre tal o cual cuestión, no obstante la evidencia en contra, o podemos tener fe en que debería hacerse tal o cual cosa, a pesar de que las acciones implicadas se pudieran considerar erróneas. Desde luego que estos tipos de fe se encuentran juntos: cuando, por ejemplo, Lenin pidió a los bolcheviques antes de la Revolución que tuvieran fe en él aunque muchas de sus acciones parecieran inexplicables, les estaba pidiendo que creyeran, entre otras cosas, que las metas del partido serían alcanzadas por sus políticas, aunque con frecuencia aparentaran ir en la dirección opuesta, y un miembro del partido tenía que continuar creyendo que el estado bolchevique era lo mejor, a lo que debería aspirarse, aunque implicara asesinatos y miseria. Se puede hacer un paralelo entre este tipo de fe y las creencias religiosas, pero en los primeros casos hay algo al menos que parece claro: *qué* es aquello que está siendo creído, pues si un hombre tenía fe en Lenin como líder del partido, o en que sus políticas

promoverían la revolución, sabía que eso era en lo que estaba creyendo, aunque no tuviera fundamentos racionales para eso. Pero es una petición extraña que alguien crea por fe algo que no entiende propiamente, pues ¿qué es lo que se le está pidiendo que crea? La fe podría ser una manera de creer en algo, opuesto a creerlo por evidencia. Pero ¿cómo podría darse el paso entre lo que se entiende y lo que no se entiende?

Puede decirse que la fe es un modo de llegar a entender algo, y aquí podría sugerirse que tiene una analogía con las artes. “Tú piensas que todo esto no tiene sentido”, podría decir alguien acerca de un poema; “pero simplemente piensa que el poeta no está tratando de engañarte, tómalo en serio, y verás de qué se trata”. El himnologista del siglo XVIII, en términos ligeramente más utilitaristas, quiso decir algo así cuando escribió: “Pon a prueba su amor; la experiencia decidirá cuán benditos son ellos, y solamente ellos, quienes en su fortaleza confían”. Pero la analogía no es lo suficientemente buena, pues nuevamente aquí la fe inicial es comprensible: la creencia de que el poema tiene un significado, que uno solamente debería encontrar. Pero en el caso de la creencia religiosa, lo que se tiene es precisamente la creencia misma, y no la suposición previa de que dicha creencia sea comprensible; suposición que puede adoptarse por fe, con la esperanza de que posteriormente se clarificará lo que dicha creencia significa. Aquí nuevamente encuentro la misma dificultad: pues si no se sabe lo que se está creyendo por fe, ¿cómo se puede estar seguro de que se está creyendo algo? Y *a fortiori*: ¿cómo puede tal creencia ser el medio para otra cosa, a saber, llegar a entender?

De cualquier modo, esto no viene al caso, pues el argumento original era que ciertas creencias religiosas deben ser inherentemente misteriosas y seguir siéndolo, y que es parte de la fe aceptarlas. Mi dificultad es que, si la creencia es incomprensible, y necesariamente lo es, uno no puede ver lo que está siendo aceptado, ni por fe ni de otra manera.

San Pablo (I Co. 1:20, 21 y 25) escribe: “¿Dónde está el sabio? ¿Dónde el docto? ¿Dónde el sofista de este mundo? ¿Acaso no entonteció Dios la sabiduría del mundo? De hecho, como el mundo mediante su propia sabiduría no conoció a Dios en su divina sabiduría, quiso Dios salvar a los creyentes mediante la necedad de la predicación. [...] Porque la necedad divina es más sabia que la sabiduría de los hombres”. Un comentarista francés, F. Godet, explica al respecto: *l'évangile n'est pas une sagesse, c'est un salut* (“el evangelio no es un sistema intelectual, sino una salvación”). Podría objetarse que mi argumento ha estado tratando las creencias cristianas demasiado como una *sagesse*, y que no se trata de un sistema de creencias coherentes y comprensibles. Esto podría formularse de otra forma diciendo que en la última parte de este artículo he pasado por alto lo que enfatice en la primera; que el lenguaje religioso no es usado solamente para hacer proposiciones, sino también para muchos otros propósitos; que las proposiciones de la creencia religiosa han de ser entendidas sólo como parte de un modo de vida, que incluye oración, observancia religiosa, etc., y podría decirse, en conexión con la discusión previa, que lo que uno escoge, cuando uno escoge creer, es vivir de cierto modo, en el cual las proposiciones desempeñan un papel. Esto es cierto, pero para que las proposiciones desempeñen dicho papel, las creencias deben estar allí, y ése es el

punto. Podemos considerar nuevamente la contracción posible del lenguaje religioso, la disminución de su alcance. Se puede dejar de mencionar a Dios en explicaciones de eventos físicos particulares, por ejemplo, o en el discurso moral, pero éstos continuarán por sí mismos como formas de discurso. Lo que no tendría sentido sería que se dejara de mencionar a Dios en las formas de observancia religiosa o en la oración, pues entonces ya no existirían en absoluto. La observancia religiosa y la oración no representan nada, hasta donde puedo ver, a menos que haya también detrás de ellas algunas creencias acerca de Dios, algunas proposiciones acerca de él. Esto, desde luego, sería el punto final de la aguja de la fe; orar precisamente al Dios no conocido, ignorando por completo si tal actividad tiene cualquier sentido en relación con él o no —o más bien, en tal ignorancia, uno tendría que decir “ello” más que “él”, y ¿podría uno siquiera decir eso? Algo debe ser creído, si es que las actividades religiosas no son simplemente silbar en la oscuridad sin siquiera el conocimiento de que lo que uno está silbando es una melodía, y algo que conecta a Dios con el mundo de los hombres. Pero tal conexión debe implicar una proposición acerca de Dios, y específicamente una que no sea explicada en términos de otras proposiciones acerca de Dios, sino en términos de la vida de los hombres. Si se dice esto, parece que debe ser o bien como una proposición no religiosa, como en nuestro burdo ejemplo pseudocientífico de la pérdida de las cosechas —que pueda hacer falsificable la creencia religiosa central y de ninguna manera lo que era requerido—, o debe ser una proposición que hable lo suficiente acerca de Dios para ser, como quien dice, misteriosa, implicando un intento de expresar la aparición de la perfección infinita de Dios en una situación humana, u otra conexión con ella. Si es un misterio en sí mismo, no puede ser explicado por la razón, pero decir que ha de ser creído por fe, y no por la razón, no encara la dificultad: pues la cuestión no era cómo debería de ser creído, sino qué era lo que iba a ser creído. Si entonces la fe cristiana es verdadera, debe ser parcialmente incomprensible, pero si es parcialmente incomprensible, es difícil ver qué significa que sea verdadera.

8

Esto es solamente la paradoja de Tertuliano con una transformación: *credibile est quia ineptum; et quia ineptum, non credibile*. Más aún, de esto se deduce que es difícil caracterizar la diferencia entre creencia e incredulidad. Podemos, desde luego, señalar que el creyente *dice* ciertas cosas que el incrédulo no dice; sin embargo, no queremos nada más esto, sino saber *qué* es lo que el creyente cree y el incrédulo no cree, pero esto es lo que no podemos hacer propiamente. Y si no podemos describir adecuadamente la diferencia entre creencia e incredulidad, no podemos ser capaces de describir la diferencia entre ortodoxia y herejía, pues la diferencia entre personas que creen diversas incongruencias es tan oscura como la que hay entre los que creen en una ineptitud y los que no creen en ella. Tertuliano, como mencioné al principio, se convirtió en un hereje.

II. Argumentos metafísicos

Los metafísicos no afirman sus posiciones. Intentan apoyarlas mediante argumentos y dar demostraciones de sus conclusiones. Cualquier investigación sobre la naturaleza de la metafísica debe contener alguna consideración sobre estas demostraciones, pues es el intento de probar sus conclusiones, de mostrar mediante un argumento lógico que tal o cual cosa debe ser cierta. Eso es lo que distingue principalmente al metafísico filosófico del místico, el moralista y otros que expresan o tratan de expresar una concepción comprensiva de cómo las cosas son o deberían ser.

Podría muy bien ser que el metafísico cabal no acostumbrara empezar hablando psicológicamente, con sus demostraciones; puede empezar más bien con una concepción del mundo y encontrar subsecuentemente demostraciones que articulan sus pensamientos en la forma requerida. En este sentido, los argumentos que da pueden ser descritos como racionalizaciones. No se entienda con esto que dichos argumentos deban ser descartados sumariamente por carecer de bases, ser inválidos o desdeñables. Parte de la palabra “racionalización” es después de todo la palabra “racional”, y es por virtud de su estructura lógica, de sus pretensiones de validez lógica, que las teorías metafísicas se distinguen de compenetraciones con la realidad simplemente intuitivas y no formuladas.

Sin embargo, la semejanza entre las teorías metafísicas y las racionalizaciones en el sentido psicoanalítico es más profunda, y uno podría preguntarse si los argumentos que el metafísico presenta realmente *importan*. Para algunos escritores recientes, las teorías y los argumentos metafísicos parecen ser simplemente síntomas de una especie de neurosis intelectual o “espasmo mental”. [28] El metafísico es un hombre con una *idée fixe* que proyecta sobre el mundo en la forma de una teoría ambiciosa y distorsionada. Así que, como no tiene caso razonar con un neurótico, tampoco lo tiene argüir con un metafísico; lo que uno debe hacer, en ambos casos, es curarlos. De ahí que vaya con esta concepción de la metafísica una correspondiente concepción del deber propio de la filosofía. El filósofo debería hacer las veces de psicoanalista con el metafísico torturado y dominado por la teoría y, pasando por alto los argumentos mediante los que racionaliza sus preocupaciones, usar técnicas analíticas para llegar a las raíces de sus preocupaciones.

En su forma más extrema, esta concepción parece ser una exageración bárbara. Lo que correctamente enfatiza es que muchos argumentos metafísicos importantes no son el tipo de argumentos que pueden simplemente ser aceptados como válidos o rechazados como inválidos mediante reglas ciertas y sobre las que hay un acuerdo general, y que su

valor o sus fallas probablemente yacen más profundamente, en algún concepto o idea central que el metafísico está tratando de articular a través de aquéllas. La debilidad de esta concepción terapéutica, si se lleva muy lejos, es que no parece haber ninguna razón por la que no debería llevarse hasta el final del camino, de manera que la metafísica llegue a ser considerada no solamente neurótica, sino como, sin duda, un tipo particular de neurosis. Si fuera así considerada, desde luego, el uso de la filosofía o del análisis filosófico para curarla sería una frivolidad; lo que se necesitaría sería un psicoanálisis real. Uno no puede creer seriamente que la metafísica sea tan irracional, o que los filósofos deberían realmente encomendársela a los médicos. El análisis que se requiere es filosófico, no psicológico, y lo que requiere no es al metafísico en sí sino sus argumentos, los cuales deberían tomarse seriamente como intentos racionales por demostrar un punto de vista.

Desde luego, los metafísicos varían en la medida en que tratan de demostrar lo que dicen; en la precisión y el rigor de dichas demostraciones, y en que éstas, no importa cómo se expresen, forman una parte esencial de su pensamiento, y no están allí solamente como decoración. Para comparar dos metafísicos británicos, por ejemplo, hay una diferencia marcada entre McTaggart y Bradley: mientras que McTaggart busca derribar formas de demostración y trozos de argumento duros como el carbón, Bradley tentativamente los esboza. No obstante, dando cabida a todas estas diferencias, hay prácticamente en cualquier metafísico occidental de importancia un núcleo de argumento; un intento de apoyar su posición o plantear sus preguntas mediante un movimiento de las premisas a la conclusión.

Todos los teóricos emplean argumentos y hacen inferencias, pues todos se dedican a llegar de un lugar a otro; de un conjunto de premisas o colección de hechos a una conclusión. Pero, lo que es igualmente notorio, no todos los teóricos hacen el mismo tipo de inferencias, y un movimiento de las premisas a la conclusión se puede hacer de acuerdo con muy diferentes tipos de reglas. Los lógicos que se han ocupado en examinar, clasificar y formalizar los diferentes tipos de inferencia, los han dividido básicamente en dos clases ampliamente distintas: inferencias inductivas y deductivas. Las inferencias deductivas son aquellas en las cuales si se aceptan las premisas, debe aceptarse la conclusión; si no, se cae en la contradicción. Se llega a la conclusión, con rigurosa necesidad lógica, de lo que implica. Tales son los argumentos, por ejemplo, de los matemáticos. Los argumentos inductivos, por otra parte, no tienen tal rigor absoluto; quien acepta las premisas no se *contradiría* a sí mismo si se rehusara a aceptar la conclusión, aunque podría verse bastante tonto. La mayoría de las inferencias prácticas de la vida cotidiana son de este tipo; así, si un hombre alcanza a partir de la experiencia personal la conclusión de que siempre es poco sabio jugar cartas con extranjeros en trenes, está dando un paso inductivo.

Este ejemplo muestra una característica adicional importante de los argumentos inductivos. El hombre que llega a esta conclusión está *generalizando*; está diciendo: “Siete veces se me ha pedido que juegue, y siete veces han tratado de estafarme, así que probablemente la siguiente ocasión, o cualquier otra ocasión, que se me pida, pasará lo

mismo”. Su argumento hasta aquí es muy válido, pero desde luego puede equivocarse en su conclusión; siempre es *posible* que el siguiente extranjero jugador de cartas pueda ser honesto. Ésta es una característica general de los argumentos inductivos. Justamente porque su función es conducir de algún hecho ya aceptado a alguna aserción más amplia o al menos diferente, sus conclusiones siempre son, en un sentido, sólo probables. Una inferencia inductiva es empírica, y siempre es concebible que su conclusión resulte ser falsa, no importa cuán cuidadosamente se haya pensado.

En el caso de los argumentos deductivos, la situación es más complicada: pueden tener conclusiones necesarias y ciertas (como las conclusiones de los argumentos matemáticos), pero solamente si las premisas son también ciertas, precisamente porque un argumento deductivo no da en la conclusión más de lo que está ya dicho o implicado en las premisas. Si las premisas son solamente probables, así también lo será la conclusión. Sin embargo, el punto que nos interesa en la discusión presente es que los argumentos inductivos sólo pueden conducir a resultados empíricos y probables.

¿Cuál es la relación de los argumentos metafísicos con estos dos tipos de argumento? Una investigación sobre esta relación debería al menos ayudarnos a ver lo que no es un argumento metafísico, y puede ayudarnos a ver algo de lo que sí es. Está claro, en primer lugar, que los metafísicos no suelen hacer inferencias inductivas directas: no hablan con la forma “tal cosa es verdadera en estos casos, así que es probable que siempre sea verdadera”. Sería absurdo, por ejemplo, suponer que un metafísico llegara a la conclusión “los hombres tienen libre albedrío” mediante un argumento como “todos los hombres que hemos observado tienen libre albedrío, así que los hombres en general lo tienen”, o bien, “todos los hombres que hemos observado tienen cejas, así que en general los hombres las tienen”. La cualidad *a priori* de una conclusión metafísica, su necesidad, por sí misma hace que tal procedimiento sea inapropiado: podría no haber necesidad de *ese* tipo de apoyo en la experiencia.

No obstante, debe decirse que algunos argumentos que los metafísicos han empleado son notablemente semejantes a las inferencias inductivas. Tal, por ejemplo, es la forma más simple del argumento teológico a partir del diseño, alguna vez conocido bajo el nombre de “el reloj de Paley”. La forma de argumento de Paley era precisamente ésta: “Si encontramos por casualidad un reloj u otra pieza de mecanismo intrincado, deberíamos inferir que ha sido hecha por alguien, pero alrededor de nosotros encontramos intrincadas piezas de mecanismos naturales, y se ve que los procesos del universo se mueven juntos en relaciones complejas; deberíamos, por lo tanto, inferir que éstas también tienen un hacedor”. Hay algunas dificultades generales acerca de un argumento analógico de este tipo, pero lo que aquí nos interesa es que parece ser precisamente un tipo de argumento inductivo. El razonamiento de Paley es simplemente éste: “siempre que en el pasado hemos encontrado mecanismos intrincados, hemos encontrado un hacedor, así que en este caso también podemos inferir uno”. Pero al ser un argumento inductivo parece demasiado débil para su propósito, pues, tomado por sí mismo, sólo puede conducir a un tipo inductivo de conclusión. Así, la proposición de la existencia de un Dios, a la cual supuestamente conduce, tendrá solamente el estatus de

una hipótesis cuasicientífica, y para cualquier hipótesis inductiva, como hemos visto, la opuesta no implica contradicción y es lógicamente posible. Así que alguien que desde *este* tipo de argumento aceptara la existencia de Dios tendría que admitir que al menos es posible que Dios no existiese después de todo, pero en general alguien que cree en Dios no admitiría que fuese en modo alguno posible que Dios no existiera; insistiría en que la proposición de la existencia de Dios debe tener algún tipo de necesidad absoluta. Así que el reloj de Paley, si es que ha de ser llamado un argumento metafísico, no parece ser característico de esta disciplina, ni siquiera un argumento característico para mostrar la existencia de Dios. Solamente es, como quien dice, una inferencia pseudocientífica.

Otra manera de decir que el metafísico no usa argumentos inductivos es señalar que en las conclusiones metafísicas la noción de probabilidad no representa ningún papel. Como hemos visto, es una característica de las inferencias inductivas que la palabra “probablemente” siempre puede deslizarse en ellas, pero el metafísico no ofrece sus conclusiones como *probables*; argumenta que *deben* ser así. Aquí nuevamente encontramos lo que parecen excepciones a esta regla. En los primeros escritos de Bertrand Russell, por ejemplo, encontramos cosas como ésta: “hay argumentos que tratan de mostrar que el mundo externo no existe, pero puesto que estos argumentos no son concluyentes, y tenemos una tendencia natural a creer que el mundo externo existe, probablemente estemos más seguros al seguir pensando que existe”. El profesor Broad y otros han escrito en una vena similar, como si sopesaran la probabilidad de una tesis metafísica en contra de otra.

Me gustaría señalar dos puntos acerca de este tipo de ejemplo. Primero, pienso que de hecho es, o ha sido, un modo de hacer filosofía que trató de asimilarla a las ciencias naturales y, por ende, consideraba sus conclusiones a la luz de probabilidades. Algunos filósofos de este siglo, impresionados por los logros de las ciencias y abatidos por la metafísica tradicional, buscaban aplicar métodos científicos dentro de la misma filosofía, y entre los resultados de eso hubo lo que parecen ser conclusiones metafísicas presentadas con un aire de probabilidad inductiva. Sin embargo, pienso que si se examinan cuidadosamente estos argumentos, se puede encontrar que no son argumentos inductivos directos, como quizá era el de Paley; el lenguaje de teoría y probabilidad es algo más que un vestido para un argumento conceptual filosófico. Más aún, estos filósofos no son metafísicos típicos; ellos mismos, sospecho, podrían haber negado que lo que estaban haciendo era metafísica.

Segundo, puede haber otra razón, muy diferente de la última, por la que Russell, por ejemplo, debería calificar sus conclusiones metafísicas con términos como “probablemente”. Ya hemos visto una razón por la que la conclusión de un argumento debería ser pensada como solamente probable: esto es, precisamente, que el argumento es inductivo, y un filósofo puede querer subrayar la naturaleza empírica de cualquier argumento inductivo insertando “probablemente” en él. Si esto es lo que quería decir Russell, entonces seguramente su argumento es inductivo, pero esto es solamente una razón por la que la palabra “probablemente” debería aparecer en un argumento. Otra razón, y muy familiar, es que las premisas del argumento son en sí mismas sólo

probables, en el sentido muy ordinario de que, por ejemplo, es (en el mejor de los casos) sólo probable que el favorito gane la carrera de las 2:30 de la tarde de mañana. Desde luego, si mis premisas son sólo probables, entonces cualquier conclusión que extraiga de ellas, incluso en un argumento deductivo, será solamente probable. Como observamos antes, no se obtiene más de la deducción de lo que se pone al principio. Así, si alguien piensa que el favorito probablemente gane la carrera de las 2:30, puede inferir válidamente, aunque sin interés, que probablemente ningún otro caballo lo hará; pero no obtendrá ninguna conclusión que sea en sí misma cierta, puesto que sus premisas no son ya más que probables. Aquí tenemos una razón por la que la palabra “probablemente” debería formar parte, incluso, de argumentos deductivos. Pero, nuevamente, volvemos al argumento de Russell, no es de *este* tipo; sus premisas —*e.g.* que creemos en un mundo externo— no son discutibles en sí mismas, como lo son las predicciones acerca de las carreras de mañana.

Hay aún otra manera en que “probablemente” puede entrar en un argumento, y esto puede arrojar más luz sobre el de Russell. Al considerar un argumento, podemos estar ocupados no tanto con la cuestión de si las premisas y la conclusión son verdaderas o falsas, ciertas o probables, como con la cuestión de si la conclusión se deduce de las premisas —esto es, simplemente podemos querer saber si el argumento es *válido*. Hablando en forma estricta, no hay grado de validez: la conclusión o se deduce de las premisas o no, y no hay punto medio. No tiene sentido decir que se llega “más o menos” o “casi” a una conclusión. No obstante, uno frecuentemente se encuentra ante la situación en que no está *seguro* de si un argumento dado es válido. Las premisas pueden ser complicadas o estar expresadas de manera no clara, o la cadena de razonamiento puede ser sutil, etc., y uno puede encontrarse ante una duda genuina de si la conclusión se cumple o no. En tales casos, uno puede expresar sus propias dudas diciendo que la conclusión “quizá” o “probablemente” se deduce de las premisas.

Este tipo de duda, y por ende este tipo de “probablemente”, puede aparecer, por supuesto, con cualquier argumento inductivo o deductivo: un problema de matemáticas, por ejemplo, puede ser tan complicado y tan poco autoevidente que lo mejor que uno puede decir, dejando pendiente mucha investigación ulterior, es que el argumento que lleva a su resolución probablemente sea válido. El argumento de Russell, y algunos otros argumentos metafísicos que conciernen a la noción de probabilidad, pueden ser de este último tipo. La noción de probabilidad entra en ellos no porque el filósofo piense que su conclusión o sus premisas sean en sí mismas cuestionables, sino porque tiene dudas acerca de la conexión entre ellas; de hecho, no está seguro de si su conclusión se cumple. Tales dudas, como hemos visto, pueden surgir con cualquier tipo de argumento, y el hecho de que un filósofo sólo se comprometa a decir que su conclusión *quizá* sea válida, no muestra en sí mismo a qué tipo de conclusión está llegando, o mediante qué tipo de argumento.

En cualquier caso, los argumentos metafísicos no parecen ser característicamente inductivos. ¿Son, entonces, deductivos?

La deducción parece un candidato mejor como herramienta profesional del

metafísico, pues los argumentos deductivos pueden al menos conducir a conclusiones de necesidad, que son lo que él quiere. Se dice comúnmente que los metafísicos buscan deducir la naturaleza de la realidad o algo por el estilo, y así se puede dar la impresión de que la del metafísico es una empresa completamente deductiva. Esta impresión parece obtener apoyo de los grandes sistemas que algunos metafísicos han construido, que pretende mostrar relaciones deductivas entre características de la realidad.

Pero la idea de que la actividad metafísica consiste simplemente en hacer deducciones en un sistema deja de lado una cuestión más fundamental. Todo movimiento deductivo debe hacerse de un lugar a otro: uno necesita tanto premisas como conclusión. Así que, en una cadena o sistema de deducciones, debe haber algo al principio a partir de lo cual comience toda una serie de razonamientos. En un sistema lógico formal, lo que se tiene al principio son axiomas; estos son, en relación con ese sistema, incuestionables. No se derivan en sí mismos del sistema; no hay nada de lo cual derivarlos. Es posible tener un número de sistemas puramente formales, cada uno con sus propios axiomas, y para propósitos particulares se puede elegir el que se desee. Pero la intención del metafísico no es darnos a elegir; él quiere hacer una serie de afirmaciones que tengan contenido y sean necesariamente verdaderas. Pero si sus conclusiones son, como él lo desea, ineludibles, y está deduciendo las conclusiones en última instancia a partir de axiomas, entonces los axiomas deben ser también ineludibles; las inferencias, a diferencia de los buzos, no ganan peso conforme se acercan al fondo. Pero los axiomas no pueden en sí mismos ser demostrados en el sistema, así que el metafísico debe tener algún otro método para apoyar sus axiomas, fuera del sistema. Tratará de mostrar que uno tiene que aceptar sus axiomas, pues solamente así puede él mostrar que uno tiene que aceptar sus conclusiones. El arma que usa para tratar de hacerle a uno aceptar los axiomas es en el sentido más fuerte el argumento metafísico.

Los constructores de los sistemas racionalistas del siglo XVII trataban a su manera de encontrar para sus sistemas axiomas ineludibles; pero su método siempre consistía en buscar axiomas que no necesitaran apoyo de ningún tipo, que fuesen verdaderos de manera autoevidente. Así que su argumento metafísico, en este punto, es más bien una apelación a proposiciones que no necesitan argumentarse en absoluto. Así, Descartes, por ejemplo, mediante su procedimiento de duda sistemática, eliminó las verdades en las que creía hasta que llegó a la verdad aparentemente indubitable “estoy dudando”, a partir de la cual dio un paso inmediato a “existo”. No me resulta claro si consideró que este paso era puramente deductivo; en cualquier caso, le pareció que la seguridad de “estoy dudando” quedaba establecida por la pura luz de la razón.

Pero no todos los constructores de sistemas usan métodos tan simples como éste para proveer sus axiomas, y no todos los metafísicos son, en este sentido más ambicioso, constructores de sistemas. Nuestro argumento ha mostrado por qué no hay necesidad de que lo sean, pues si un argumento metafísico puede ser usado para que se acepte una proposición que ha de ser entonces utilizada como axioma de un sistema, también puede ser usada para que se acepte la proposición incluso si no ha de ser utilizada como un axioma. Incluso los metafísicos que no están relacionados con la construcción de

sistemas comprensivos tratarán de mostrar mediante argumentos constructivos que tal cosa (que por lo demás uno no esperaba) debe ser así; o muy frecuentemente, mediante un argumento destructivo, que tal otra (que uno sí esperaba) no puede serlo. Trataré de explicar en un momento algo acerca de la interrelación de esta clase de argumentos en un caso típico.

Al contrastar hasta este punto los argumentos usados por metafísicos con argumentos inductivos y deductivos, he hablado como si hubiese un tipo definido de argumento que fuera metafísico. Pienso que, de hecho, no hay una cosa única que sea un argumento metafísico, así como tampoco hay una cosa única que sea una proposición metafísica. Esto no significa, sin embargo, que no se puedan hacer observaciones absolutamente generales acerca de ellos; más bien tales observaciones generales servirán para caracterizar estos argumentos en bosquejo más que para formular alguna propiedad esencial de ellos. Una propiedad esencial de ellos, sin embargo, puede ser y ha sido formulada: que no son lo mismo que argumentos deductivos o inductivos. Para algunos filósofos, desde luego, como los así llamados positivistas lógicos, esto es más que suficiente; todos los argumentos y proposiciones metafísicos son amontonados por ellos y desechados como carentes de sentido, pero esta unidad esquelética no nos dice nada acerca de la metafísica; es solamente la uniformación de todo. El enfoque de estos filósofos colectivamente antimetafísicos es una especie de paralelo filosófico a la actitud de una persona fanáticamente militarista que divide a todos los hombres solamente en dos clases, combatientes y no combatientes. Pero incluso desde el punto de vista militar tal división tendría sus desventajas: entre los no combatientes, por ejemplo, deja de distinguir entre los que están médicamente incapacitados y los objetores de conciencia. Un metafísico, en relación con el criterio positivista de significado, es más como un objetor de conciencia que como un inválido; es su entero propósito hacer algo diferente de lo que el positivista quiere que haga. Y justamente como hay diferentes tipos de objetores de conciencia, así también hay diferentes tipos de metafísicos y de argumentos metafísicos. Sólo se puede adquirir entendimiento considerando los casos individuales.

Dentro de los límites de este capítulo sólo es posible abordar con mayor detalle un ejemplo de argumento metafísico. Tiene la abrumadora desventaja de ser sólo un ejemplo, pero es tan central y recurrente, que espero pueda arrojar también algunas conclusiones de validez general.

El escenario para el argumento que voy a considerar se basa en los hechos de la ilusión de percepción. Alrededor de nosotros reconocemos objetos de diversos tipos — árboles, mesas, personas, etc.—. Ocasionalmente, cuando se trata de reconocer cosas, nos engañamos, y damos por hecho que algo que vemos es determinada cosa que en realidad no lo es. Así, una vieja bota en la penumbra se puede confundir con un gato pequeño. En este escenario aparece el filósofo. Puede plantearse un número de preguntas, como “¿qué es lo que realmente sabemos?”; “¿cuánta confianza se puede poner en la percepción como fuente del conocimiento?”; “¿qué existe realmente?”, etc. Pero no importa cuál sea su pregunta particular; su razonamiento sobre la ilusión de percepción puede muy bien proceder como sigue: “Te confundiste cuando pensaste que

la bota era un gato. Puesto que te engañaste, no puede haber diferencia *intrínseca* entre la experiencia que tuviste al momento de ver lo que de hecho era una bota en la penumbra y la experiencia que pudiste haber tenido al ver realmente un gato en ese momento. La diferencia, después de todo, vino después: cuando viste más de cerca, hiciste un ruido o lo que haya sido. De hecho no puede haber diferencia intrínseca entre las dos experiencias, pues si la hubiera, habrías percibido la diferencia, y no hubieras tomado una cosa por la otra. Así que, ¿en qué consistió tu experiencia? Desde luego no fue ver un gato, pues no había ningún gato, pero la experiencia que tuviste es evidente que no fue tampoco la de ver una vieja bota. Pues bien, ya hemos estado de acuerdo en que la experiencia es similar a la de haber visto realmente un gato, pues de otra manera no hubieras podido confundir la bota con un gato, y si hubieras podido tener esta experiencia al ver realmente un gato, la experiencia no puede ser simplemente la de ver una bota; pues cuando tú realmente ves un gato, en realidad no ves una bota. Así que la experiencia que tienes en ambos casos debe ser algo neutral entre realmente ver un gato y realmente ver una bota; es algo común a ambas y menos que cualquiera de ellas”. “Más aún —este filósofo podría continuar— “el hecho de tener la experiencia visual puede ser más básico que el de ver objetos reales; pues se pueden tener experiencias visuales de hecho sin ver el objeto apropiado o incluso ninguno, pero no se puede ver un objeto sin tener una experiencia visual”.

Así procede el argumento de la ilusión en una de sus muchas formas. Contiene un movimiento destructivo y otro constructivo; ambos son típicos del argumento metafísico. El movimiento destructivo consiste en mostrar que algo que naturalmente deberíamos decir si se nos pidiese que reflexionáramos sobre la percepción —esto es, que claramente vemos objetos— es falso. Puede decirse que no hay nada demasiado sorprendente acerca de esto, y que cualquiera que dijera que nosotros siempre, cuando vemos algo, vemos un objeto material, obviamente estaría equivocado, pero que, no obstante, nosotros a veces vemos objetos materiales. Sin embargo, el movimiento destructivo es más fuerte que eso. El metafísico pretende, de hecho, que *no hay* diferencia entre ser engañado y no ser engañado; su argumento es simplemente que la diferencia no se halla donde se esperaba que estuviese, pues su argumento pretende mostrar que mediante la reflexión sobre los casos de ilusión se puede llegar a ver que los casos de percepción genuina también son diferentes de lo que pensamos; que en éstos, también, la experiencia visual del observador —que, continuará argumentando el metafísico, es privada al observador— debe desempeñar un papel.

Decir esto, sin embargo, es haber empezado ya el movimiento constructivo del argumento. Es característico de los argumentos metafísicos que el método de la destrucción señale ya lo que ha de ocupar el lugar de las cosas que han sido destruidas. Así, por eso lo que es, en un sentido, el mismo argumento —un argumento, en cualquier caso, generado por los mismos hechos de la experiencia— puede aparecer de diferentes formas en diferentes filósofos para adecuarse a sus varios propósitos. La forma en que he presentado el argumento de ilusión (y alguna forma particular tiene que elegirse) puede, de hecho, poner los fundamentos para una metafísica empirista usando la noción

de una “idea” o una “sensación”.

Pero el mismo argumento se puede usar para fines muy diferentes de aquellos del metafísico empirista. Platón, por ejemplo, aceptaba la primera etapa del argumento, y razonaba, a partir de ésta, que nuestras creencias acerca del mundo material deben ser personales, efímeras e inestables. Agregó las premisas de que el conocimiento verdadero debe ser inmutable y estable, y de que podemos hasta cierto punto tener conocimiento, y alcanzó la conclusión de que debe haber un mundo de cosas inmutables, el mundo de las formas. Lo que él y los empiristas tienen en común es el uso del argumento para destruir un mundo que se da por sentado y poner en su lugar otra cosa; en su caso, un mundo de formas; en la de ellos, una sucesión de experiencias a partir de las cuales los objetos tienen que ser inferidos o contruidos de alguna manera. Aquí vemos una característica primaria del argumento metafísico: su uso para establecer proposiciones de existencia o inexistencia. “El mundo de las formas es el mundo de la existencia genuina”; “los componentes últimos del mundo son *sense-data*”: éstas son proposiciones metafísicas (muy diferentes), y el puntal principal es el argumento que hemos estado examinando.

¿Cómo está este argumento relacionado con la distinción entre argumentos inductivos y deductivos? Es claro que no es sólo inductivo: el metafísico empirista, por ejemplo, no está diciendo precisamente “hay ilusiones, así que probablemente los objetos no existen y las experiencias individuales sí”. No obstante, está haciendo un movimiento que trasciende aquello con lo que empezó. Al examinar el concepto de ilusión de percepción, llega a una conclusión general acerca de la percepción como tal; una conclusión que ha de ser atacada, no mediante la producción de cualesquiera contraejemplos directamente empíricos, sino mediante una investigación de sus conceptos, en particular el más bien dudoso concepto de una “experiencia”.

“Bien —alguien podría decir—, todo lo que tenemos en este argumento es un hecho contingente y un conjunto de deducciones. Las deducciones se hacen a partir del concepto de una ilusión de percepción; el hecho contingente es que el concepto tiene aplicación; esto es, a veces resultamos engañados”. Pero esto sería un completo malentendido, pues ni es el hecho supuestamente contingente sólo un hecho contingente, ni son las supuestas deducciones sólo deducciones. Cuando el metafísico dice, con un solapado aire de simplicidad fáctica, “a veces somos engañados, y tomamos una cosa por la otra”, no está simplemente formulando un hecho fortuito, algo que podría ser de otra manera. Todo lo que realmente necesita para su argumento es la posibilidad lógica de la falla en el reconocimiento, la existencia de tal concepto, y, muy a grandes rasgos, el concepto del reconocimiento se halla contenido en la posibilidad de la falla en el reconocimiento. En tanto que tenemos el concepto de reconocer cosas, también tenemos el de fallar al reconocerlas. Desde luego, *quizá* podríamos no tener ninguno de esos conceptos, pero qué parecería ser entonces nuestra percepción, es totalmente oscuro. En la relación de reconocimiento y conceptos similares a nuestra experiencia se halla un problema filosófico enorme. Nuevamente, las deducciones no son simplemente deducciones. Si lo fueran, apenas podría existir el desacuerdo que hay con la conclusión, y nuevamente el metafísico ha adquirido *en el camino* un concepto con el cual no

comenzó, el de una “experiencia”.

No obstante, la introducción de este concepto no es simplemente gratuita. Parece ser algo implicado en lo que ya se ha dicho, ser exigido por los hechos tal y como están, y una meta principal del argumento metafísico es exhibir los hechos para mostrar de dónde proviene tal exigencia. El propósito del argumento no es sólo deducir una conclusión a partir de los hechos; es más bien mostrar que la explicación de aquellos hechos, cuando reflexionamos sobre ellos, tiene un agujero en ella, un agujero que es en el que encaja exactamente el especial concepto del metafísico. Este concepto puede ser, como el de “experiencia”, que exista ya en una forma ordinaria en nuestro lenguaje, y que el metafísico recoge, dignifica y refina para convertirlo en un principio de explicación. Por otro lado, si es un metafísico muy acucioso, los conceptos que usa de este modo pueden ser mucho más técnicos y alejados del pensamiento ordinario, como las “mónadas” de Leibniz o “los objetos nouménicos” de Kant.

La grandeza de un metafísico, me parece, ha de determinarse por tres consideraciones: cuán arbitrarios son sus conceptos especiales, cuánto explican y cuánto distorsionan nuestro pensamiento ordinario. Estas consideraciones no son, desde luego, independientes; están unidas, como los requerimientos del diseño de un aeroplano en los que los diseñadores expertos tienen que conciliar exigencias en conflicto como peso mínimo, capacidad máxima y requerimientos de seguridad. El diseñador de genio se acerca tanto como puede a tener el mejor de los mundos posibles, y así lo hace el genio metafísico. Sus conceptos explicarán mucho, revelando analogías importantes entre tipos de experiencia y pensamiento que superficialmente parecen ampliamente diferentes. Estas analogías deben ser reales, y no el producto de asimilaciones forzadas o en exceso distorsionantes, y no deben ser arbitrarias, en el sentido ulterior de que uno debe ser conducido a reconocerlas, y con ellas también la exigencia de los conceptos explicativos del metafísico, mediante un argumento claro y lógico.

Pero es el argumento lo que nos ocupa aquí, más que las características de la empresa metafísica cuando ha sido completada. Cualquier explicación de tales argumentos, en unas cuantas palabras, está destinada a ser una caricatura, pero sus características estándar se pueden resumir como sigue: el metafísico siente inconsistencia, dificultad o incompletud en lo que naturalmente tendemos a pensar acerca de alguna de las características de nuestra experiencia, o más bien en lo que parece ser presupuesto en lo que así pensamos. Al resolver esto, tratará de mostrar que algún concepto en el cual confiamos es secundario para, o está presupuesto por, algún otro concepto que ha introducido o extendido desde otra parte. Este concepto suyo puede tener un lugar especial en la respuesta al problema en cuestión (como el uso de “experiencia” en los empiristas), o puede usarlo ampliamente en otra parte (como las formas de Platón) para resolver otros problemas; cuanto más amplios sean los usos que le dé en otro lugar, más sistemática será su filosofía.

La compulsividad de su argumento provendrá de que empieza con conceptos y características de la experiencia que, parece, deben encontrarse allí si hemos de pensar acerca de nuestra experiencia en absoluto. Su intento de mostrar que algún concepto

implica una dificultad, o está presupuesto por algún otro concepto, frecuentemente resultará en aserciones de existencia o no existencia. No obstante, éstas, a diferencia de las aserciones de existencia empírica o matemática, son en un sentido solamente comparativas, pues todos los metafísicos están de acuerdo en que la apariencia (aquellas características del mundo que se muestra metafísicamente que no son reales) debe finalmente encontrar *algún* lugar en la explicación de las cosas como realmente son. Vimos esto antes en la preservación empírica, en un lugar diferente, de la distinción ordinaria entre ilusión y percepción genuina. Incluso, la famosa demostración de McTaggart de la irrealidad del tiempo (que es tanto filosóficamente espectacular como muy difícil de refutar), prepara una explicación de qué es lo que realmente existe y se presenta a nosotros confusamente como el paso del tiempo. De allí que algunos hayan visto la actividad del metafísico primariamente como una reubicación: la extensión de algún concepto favorecido a un lugar primario en la explicación de las cosas a expensas de los conceptos más familiares.

Hay verdad en esto; no obstante, la elección de tal concepto y el punto de su aplicación no es simplemente arbitrario, y se nos deja con el problema de por qué algunos funcionan mejor que otros. Los argumentos metafísicos son como árboles; su posición exacta o su configuración son hasta cierto punto cuestión de preferencia: el metafísico puede escoger dónde exactamente plantarlos y cómo podarlos, pero no puede escoger si crecerán o no; algunos puntos del paisaje conceptual son más fértiles que otros. Si con el hacha positivista derribamos los árboles, vuelven a crecer; si con la pala wittgensteiniana empezamos a desarraigarlos, proseguiremos fascinados sin cesar. Pues, incluso, si extraemos algún conjunto de las raíces, habrá, si es un árbol robusto, muchas otras. Quizá excavar sea la actividad filosófica propia en este tiempo; de hecho el mero orgullo por haber hecho crecer un árbol más grande que el de cualquier otro ya no es suficiente. Pero había algo que justificaba tal orgullo: el conocimiento de que el metafísico, con su habilidad para la jardinería, había encontrado el sitio donde las bellotas podían crecer. Lo que motiva al excavador filosófico es el deseo por saber más. ¿Qué hace crecer los árboles metafísicos? ¿Por qué los argumentos metafísicos surgen de algunos rasgos de nuestra experiencia más que de otros? La respuesta a esa pregunta sería la última respuesta metafísica.

III. Placer y creencia

Podemos estar de igual manera complacidos por algo que solamente creemos que es así pero no lo es, como por lo que sabemos que así es. De esta manera, puedo sentir agrado porque (como lo supongo) he heredado una fortuna, cuando esto no ha sucedido. Este hecho merece consideración, en particular porque plantea la pregunta de la relación entre el placer y sus objetos; este artículo se ocupará principalmente de esta cuestión.

Si alguien está tentado a pensar que el objeto de mi placer —lo que me produce agrado— es la *causa* de mi placer, este tema debería desanimarlo, pues el objeto del placer, en este caso, parece ser una herencia, pero ésta, puesto que no existe, no puede ser una causa (como solía decirse, *non entium non sunt effectus*). No obstante, la causa del placer no puede ser algo muy diferente de la supuesta herencia, pues si me persuado de que de hecho no he heredado una fortuna, mi placer desaparecerá, así que debe haber estado conectado al menos con mi creencia previa en la herencia. Más aún, al hablar retrospectivamente del placer, diré que sentí agrado porque creí que había heredado una fortuna, o algo así.

Por ende, el único modo de explicar la causa del placer y sus objetos sería decir que la causa de mi placer era la creencia en la herencia. No obstante, esta explicación, a su vez, plantea dificultades. En primer lugar, si mi creencia en la herencia era la causa de mi placer, debe haberlo sido por virtud de alguna ley que conectase tales creencias con el placer, pero, ¿qué ley es ésta? Evidentemente, creer en una herencia no es la *causa* de *ningún* placer, sino, en el mejor de los casos, del *placer ante una herencia*; no obstante, es esta última noción la que se suponía que explicaría la causa. Sin embargo, se puede replicar a esto que no hay necesidad de introducir esta noción en la ley causal: será suficiente decir que la creencia en una herencia se halla entre las posibles causas del placer, si entendemos el placer mismo como un sentimiento (por ejemplo), que es el mismo sin importar la causa. Así que cuando decimos que alguien “siente agrado porque cree que ha heredado una fortuna”, lo que queremos decir es que “está experimentando placer, y la causa de esto es su creencia en que ha heredado una fortuna”.

De esta concepción se llega a que es *siempre* la creencia lo que causa el placer, incluso en aquellos casos en que la cosa que agrada realmente existe; pues si no, la proposición “siento agrado porque he heredado una fortuna” expresaría una hipótesis causal diferente de —e incompatible con— la hipótesis expresada por la proposición “siento agrado porque creo que he heredado una fortuna”. Pero es evidente que, al tiempo en que creo en la herencia, no puedo tener fundamentos cualesquiera para

preferir la segunda de estas hipótesis a la primera, puesto que me es lógicamente imposible distinguir entre lo que (como yo creo) es, y lo que yo creo que es. Por ende, habrá dos hipótesis incompatibles, que en principio no seré capaz de distinguir, acerca de mi placer, pero está claro que mi descripción retrospectiva de la situación como “sentí agrado porque creía...”, y la descripción de cualquier otra en estos términos, están precisamente basadas en mi sincero pensar o decir en el momento “siento placer porque tengo...”. Así que tal parece que una condición necesaria de la afirmación de la verdadera hipótesis sería mi creencia previa en (o la afirmación de) una aseveración falsa, y esto es absurdo.

Por ende, la explicación causal debe sostener que es siempre mi creencia la que es la causa, o al menos la causa próxima, de mi placer, y que la proposición “siento placer porque he heredado una fortuna” debe interpretarse como “siento placer porque creo que he heredado una fortuna”. Sin embargo, esto es igualmente inverosímil. En primer lugar, a partir del argumento previo, al parecer es muy poco probable que esté cerca de llegar a la hipótesis correcta y distinguirla de las rivales; por lo menos, una condición necesaria sería que me hubiese implicado en la reflexión filosófica; en segundo lugar, es imposible ver qué evidencia podría tener para la hipótesis, o cómo podría ponerme a reunir evidencias; en tercer lugar, puesto que la proposición en cuestión expresa, desde esta perspectiva, una hipótesis causal, sería corregible, y tendría sentido decir simplemente que me equivoqué al pensar que era cierta creencia la que me daba placer; pero, en general, ningún sentido se le puede atribuir a esto. De hecho, esta imposibilidad de encontrarle alguna coherencia se extiende incluso a las formulaciones que no incluyen una referencia explícita a mis creencias. Hay, desde luego, modos en los que puedo estar equivocado, o ser ignorante de los objetos de mi placer, y será una de las metas de este artículo investigarlos, pero no puedo equivocarme al decir “siento placer porque he heredado una fortuna” del mismo modo en que puedo estarlo al decir, por ejemplo, “tengo dolor de estómago porque comí alguna fruta podrida”.

Así que el objeto de mi placer —aquello que me agrada, o (en este sentido) debido a lo cual me agrada— no ha de ser tomado como una causa: ni puede mi creencia en tal o cual cosa hacer que ésta funcione como la causa de mi placer.

De hecho, la idea completa de que las creencias del hombre son una causa de placer en tales casos es una ficción ayudada, aunque no inspirada por, un malentendido en la forma “sintió placer porque creía”. Ahora bien, esta forma, y tal vez un malentendido similar, ocurre también en otra conexión: la de un hombre con fundamentos erróneos para una acción. Hay dos semejanzas obvias en los dos casos. Proposiciones de la forma “lo hizo porque creyó que p ” están basadas, en última instancia —como las proposiciones comparables acerca del placer—, en la propia proposición del hombre, tomado como sincero, de la forma “lo hice porque p ”; y las proposiciones de la última forma no son susceptibles de ser directamente erróneas. Quizá valga la pena notar que hay lenguajes que no emplean en absoluto la fórmula que conduce a error “porque creyó”, sino que ejecutan la misma función solamente por el modo de un verbo.

Así que, al menos en este sentido, hay alguna analogía entre el placer y sus objetos

por una parte, y las acciones y sus fundamentos por la otra. Pero esta analogía no nos llevará muy lejos, pues el placer, al igual que muchos otros estados con los cuales surgen dificultades similares, como el temor, no es una acción, sino más bien, al menos, una “pasión”, o algo que nos sucede; pero incluso si esto se niega, la analogía deja de funcionar. Además, si estamos de acuerdo con Aristóteles,[29] y posiblemente con el profesor Ryle,[30] en que el placer en el caso habitual consiste en (o acompaña) una actividad entusiasta, tendría que ser la actividad, al menos, lo que constituiría el objeto del placer, pues es esto *en* lo que me complazco, pero si es la actividad la que constituye el objeto del placer, no estaría basada en los fundamentos de la actividad, si es que la hay; así que, incluso en este caso, la relación entre el placer y su objeto no habrá sido explicada como la relación entre las acciones y sus fundamentos.

De hecho, en muchos casos es imposible descubrir alguna actividad que genere placer por una mera relación entusiasta en ella. El hombre que siente agrado porque cree que ha heredado una fortuna puede, desde luego, disfrutar “actividades” como imaginar su estilo de vida mejorado, planear vacaciones caras, visualizar la gratitud de las personas con las cuales será generoso, etc., pero incluso si supusiéramos, lo que parece ser falso, que tales “actividades” son el concomitante lógicamente necesario de sentir agrado ante una supuesta herencia, su realización entusiasta no puede, de hecho, constituir el placer en cuestión, pues un hombre puede disfrutar de tales actividades (como soñar despierto) incluso cuando sabe que esas cosas no llegarán a suceder. Disfrutará las actividades de modo muy diferente, y mucho más, y dará una respuesta muy distinta a la pregunta ¿por qué sientes agrado? si realmente cree que dichas cosas ocurrirán, o *porque* simplemente lo cree. Por ende, disfrutar de tales actividades no es una condición suficiente para sentir agrado porque creo que he heredado una fortuna, y no hace nada para explicar las peculiaridades de esto último. Así que la única “actividad” que nos queda como objeto de este tipo de placer es creer en estos futuros eventos, y esto no funcionará, pues atribuimos sentido a “recibe placer al creerlo” solamente donde el hombre no lo cree (realmente), o, en todo caso, donde hace un esfuerzo por creerlo, o se ha rehusado a convencerse de lo contrario, y éstos son asuntos muy diferentes. Una actividad en la que me puedo deleitar es seguramente algo en lo que me puedo conectar o dejar de lado; que puedo asumir o abandonar, y ninguna de estas cosas se pueden, en general, hacer sin convicción.[31]

En tales casos, el placer ciertamente no puede consistir en ninguna actividad entusiasta. Así, por más de una razón, la ligera analogía epistemológica entre los fundamentos de la acción y los objetos del placer no se puede buscar directamente. ¿Cómo ha de caracterizarse entonces la conexión entre el placer y sus objetos?

Consideremos la noción de “apuntar a”. Esto nos conducirá, me temo, a un gran rodeo, pero, finalmente, de regreso al placer. Es una cuestión familiar que el mero hecho de que mi dedo apunte a algo, en el sentido de que una línea dibujada desde él encontrará tal cosa, no significa que esté apuntando hacia ella.[32]

Es característico del sentido de “apuntar a”, que si se me pregunta a qué estoy apuntando, deba ser capaz de dar una respuesta. Incluso que, sujeto a ciertas reservas,

no puedo estar equivocado o ser ignorante, *i.e.*, que no tiene sentido decir que estoy en el error o que ignoro qué estoy señalando. Más aún, que puedo apuntar a cosas de diversas categorías. Consideremos lo que puede venir después de “mira” —esa cosa, su tamaño, su configuración, el número de hormigas ahí abajo, lo que él está haciendo, el color del cielo, la velocidad a la que va, etcétera. En último lugar, que al señalar puedo llamar la atención de alguien sobre alguna de esas cosas, y que el propósito de apuntar es generalmente, si no es que siempre, hacer precisamente eso.

Ahora bien, no es auténtico decir sin reservas que no puedo estar equivocado o ignorar lo que estoy señalando. Pues puedo estar equivocado, o algo anda mal con mi explicación de lo que estoy señalando, al menos en los siguientes casos:

- a) Acerca del pasado, puedo recordar que había señalado algo, pero recordar mal (o lo que es menos usual, olvidar completamente) qué era aquello que estaba señalando. Esto es obvio y carente de interés.
- b) Puedo apuntar a x , pero erróneamente decir que estoy apuntando a y , porque erróneamente creo que x es y . La primera posibilidad es que dicho error no importe, porque no era *como* y lo que estaba apuntando como x . El caso más común de esto es que se quiera apuntar a la actitud, cualidades, etc., de x , e identificar a x como y (“mira el maravilloso color de estas flores de hibisco”. “¿Qué? No hay flores de hibisco”. “¡Oh!, quiero decir el color de ese arbusto allí, sea lo que fuere”).
- c) Este caso es como *b*), excepto que aquí el error sí importa, porque estoy apuntando a x como si fuera y . (“¡Mira, la reina!”. “No es la reina”. “¡Oh!”) Aquí es verdad que apunté a algo, y algo de lo cual pude haber dado otra descripción (*e.g.*, “esa dama que acaba de entrar”), pero aquí apuntar es, como fue el caso, reservado, y la atención atraída a la cosa por mucho desaparece.
- d) Puedo apuntar a una x que no existe en absoluto, como Macbeth podía haber apuntado a la daga. Aquí la descripción dada (“una daga”) tiene que ser retirada, pero en su lugar se puede dar un tipo diferente de descripción (“estaba apuntando a la daga que pensaba que estaba ahí”). Es digno de notarse que “la daga que pensaba que estaba ahí” no es una descripción referente a una imagen, o a cualquier cosa particular semejante. Es una descripción cuya aplicación se halla en el mundo externo, pues la daga que pensaba que estaba ahí es (si se me permite la expresión) precisamente la que no está realmente ahí.

Hasta aquí, lo que concierne a los errores que pueden surgir con mis descripciones de lo que estoy apuntando, o a lo que estoy atrayendo la atención de alguien. Si consideramos ahora la situación correlativa de que alguien atraiga mi atención, vemos que las cosas son diferentes, puesto que, desde luego, puedo estar equivocado precisamente respecto de lo que está atrayendo mi atención. O las cosas pueden cambiar, y puedo tener una mejor idea que él acerca de las descripciones que da, *i.e.*, puedo ser capaz de corregir sus descripciones de inmediato. Pero ahora consideremos la situación en que una cosa atrae mi atención no por una persona que apunta, sino por la cosa

misma; mi atención puede ser atraída por y hacia cualquiera de los tipos de cosas a las que yo o alguien más pudiera apuntar. Ahora bien, aquí no puedo cometer el tipo de error que surge cuando alguien más apunta y yo lo malentiendo, pues eso está relacionado con la descripción que daría, y eso no sucede en el caso presente. Tampoco me puedo corregir, como puedo corregirlo a él, en el lugar; aunque puedo proceder a corregirme a mí mismo, o si comparto el objeto de atención con otros al señalar, la situación es como la descrita arriba. El problema es *más bien* con mi propio señalar, y aquí es importante que poner mi atención en una cosa es frecuentemente un paso preliminar a apuntar hacia ella, y que la descripción que daría al apuntar a ello es frecuentemente la descripción bajo la cual, por así decirlo, la cosa atrae mi atención sobre ella. Que mi atención sea atraída por algo, por supuesto, difiere de apuntar a ella, en tanto que lo segundo es una acción y la primera es más bien algo que me sucede. Esta diferencia es importante, pero lo es más que la diferencia no afecte la naturaleza de las descripciones que se dan en los dos casos. Las descripciones o identificaciones, buenas o malas, que pudiera dar de lo que atrajo mi atención, son justamente las que podría dar si fuera a llamar la atención de alguien hacia el objeto en cuestión, y justamente éstas pueden ser equivocadas en alguno de estos modos; no obstante, no tiene sentido decir que estaba en el error al suponer que estaba apuntando a esto más que a aquello, así cuando yo caracterizo lo que atrajo mi atención puedo describirlo erróneamente y, no obstante, no tiene sentido decir que estaba en el error al pensar que era esto lo que atrajo mi atención más que aquello.

Hay un peculiar caso en el que puedo “tener mi atención puesta en algo” pero no “señalarlo”, lo cual puede agregarse a la presente lista de ejemplos, aunque no es en absoluto, estrictamente hablando, un caso de error o ignorancia.

- e) Puedo poner mi atención en alguna característica que ya sé que es ilusoria, pero que no obstante es notable y digna de atención. Así, después de que he tomado mezcalina, mi atención puede ser atraída por la apariencia inusual de mi alfombra. Aquí sé que estoy bajo una ilusión, así que, desde luego, no trataré de señalar esta extraordinaria apariencia como si fuera verdad.

Un caso adicional tiene que ver con la ignorancia:

- f) Puede ser que x haya atraído mi atención, pero que no sepa exactamente qué hay en x que haya atraído mi atención.

No hay analogía para este último caso, tampoco en el apuntar. Debe haber, por supuesto, algo que haya atraído mi atención, aunque sólo sea porque me atrajo más que otra cosa, y, a partir de la descripción dada, seré capaz de señalarla, a menos que, desde luego, el caso sea peculiar en otro sentido, por ejemplo, del tipo *e*. Pero tiene sentido decir que no sé qué es lo que atrajo mi atención en ella. No hay nada sorprendente en esta diferencia: puesto que mi propósito al apuntar es llamar la atención de alguien hacia x , la descripción que doy de x supongo que logrará este propósito, y dar una descripción

indeterminada debido a la ignorancia no será apropiado. Es, desde luego, factible que apunte a x , y atraer la atención de alguien a x (con alguna descripción) justamente porque no sé qué es (con alguna otra descripción, más específica): “mira ese pájaro, ¿qué es?”, o “¡mira esa configuración en la esquina!” Esto es análogo a tener mi atención puesta en algo y no saber en qué parte de ello, pero sólo es parcialmente análogo, pues en el caso de apuntar, si yo apunto a x , y pregunto qué es, cualquier otra descripción que proceda a dar de ello en un intento por responder a la pregunta no se halla en ninguna posición lógicamente diferente de la sugerencia de alguien más, pero la situación es diferente con la cosa que atrajo mi atención. Si intento averiguar qué había en ella que atrajo mi atención, parece que, aunque otros puedan hacer sugerencias y tener teorías sobre el asunto, me corresponde a mí decidir si sus sugerencias son correctas. Si yo sinceramente y de todo corazón decido que fue una particular característica de esta cosa la que atrajo mi atención —*e.g.*, cierta similitud, aunque fuese imaginada, entre el sombrero de esta dama y un punto de referencia familiar— tiene poco sentido decir que me equivoqué en el diagnóstico. La semejanza imaginaria puede resultar ser solamente caprichosa; pero no puedo equivocarme al decir que era esta semejanza la que atrajo mi atención hacia la forma de su sombrero; o que fue la forma en sí misma la que atrajo mi atención hacia el sombrero; o que fue el sombrero lo que atrajo mi atención hacia la dama. A lo sumo, podría decir en ciertos contextos muy complicados que me estoy engañando; por ejemplo, una persona inhibida *podría* autoengañarse con éxito al pensar que lo que atrajo su atención hacia cierta muchacha era el material inusual de su vestido más que su corte provocativo, pero tales casos son quizá raros.

Es un problema confuso cómo es que uno puede llegar a saber a partir de la reflexión que era *esta* característica la que atrajo la atención de uno hacia algo. Es evidente que tales conclusiones no se alcanzan mediante inferencia empírica, y que uno no ha establecido ningún tipo ordinario de proposición causal.[33]

Antes de que regresemos a discutir directamente el placer, se debe hacer un comentario adicional acerca de “tener mi atención puesta en”. No se han distinguido dos tipos de casos significativamente diferentes. En el primero, mi atención es atraída por alguna cosa porque estoy esperando o buscando algo y dicha cosa parece un buen candidato a ser el objeto de mi expectativa o búsqueda. En tales casos, obviamente, mi atención se desvanecerá si resulta que no es el objeto en cuestión y, si lo es, resultará el objeto de un diferente tipo de atención, a saber, del que motivó la expectativa o la búsqueda. En el segundo caso, algo puede atraer mi atención por parecerme sorprendente o (más bien) intrigante. En este caso, la atención se desvanecerá si resulta que el objeto no es sorprendente en absoluto, o sí lo es, pero la sorpresa se disipa mediante la explicación; o si, sin explicarla, simplemente me aburro con ella o me distraigo con otra cosa. La diferencia más importante entre estos dos casos es que en el primero el objeto atrae mi atención porque supuestamente encaja en una descripción (en el sentido más amplio) para la que ya estoy preparado; una descripción que figuraría en una explicación de lo que ya estaba buscando, esperando, etc., y en el segundo, la descripción no está preparada; el objeto simplemente se introduce por sí mismo; llega sin

invitación. Se deduce de esta diferencia que las situaciones *a-f*, descritas arriba, no surgen de la misma manera en los dos casos; por ejemplo, *f* no surge en absoluto en el primer caso.

Regresemos ahora al caso del placer. Es obvio que hay una analogía estrecha entre los casos *a-f* en conexión con la atención, y casos similares en conexión con el placer. Hay, de hecho, un paralelo para cada caso, en las siguientes líneas:

- a) Primero, y nuevamente sin importancia, puedo recordar alguna ocasión en el pasado en la que sentí agrado, pero recordar mal u olvidar completamente qué fue lo que me agradó. Esto es quizá raro, pero lo es menos con el caso inverso del sufrimiento: puedo recordar los sufrimientos de la niñez o la adolescencia, pero olvidar sus objetos, porque ahora lo consideraría trivial. En este caso no debe considerarse la posibilidad de que el placer presente se base en recordar mal. Esto se verá posteriormente.
- b) Puede agrardarme *x*, pero digo que *y* me agrada porque equivocadamente creo que *x* es *y*, pero esto no importa porque el que *x* sea *y* no es ningún elemento de mi placer. Así que me puede agradar esta pintura y decir que me agrada este Giorgione, cuando la pintura no es un Giorgione.
- c) Más drásticamente, me puede agradar o puedo obtener placer de *x*, el cual, erróneamente, pienso que es *y*, cuando el supuesto ser *y* de *x* es la base de mi placer. Así, puedo sentir agrado por este supuesto Giorgione como si fuera un Giorgione.
- d) Me puede agradar algo que no existe en absoluto, por ejemplo, mi supuesta herencia de una fortuna.

Común a los casos *c* y *d* es la característica de que el descubrimiento de la verdad significa el fin del placer, al menos, de *ese* placer. La distinción entre el caso *c* y el caso *d* es con frecuencia sólo un asunto de expresión, y en muchos casos sería absurdamente bizantino tratar de ponerlos por la fuerza en una clase o la otra. Por ejemplo, puedo sentir mucho placer por la llegada a mi fiesta de un caballero, de quien pienso que es cierto autor distinguido. Esto podría, naturalmente, caber en la clase *c*, pero alguien muy estricto podría argumentar que aquello que me agradaba era un (supuesto) acontecimiento, a saber, la llegada de este autor, y que esto no ha sucedido, es decir, no existe, así que el caso pertenece a la clase *d*. Sería seguramente frívolo insistir en una decisión entre estas dos explicaciones.

Sin embargo, quizá emerge aquí un punto más serio: podría objetarse que la frivolidad de un argumento sobre esta última cuestión muestra una debilidad o una total falta de claridad sobre los *objetos* de placer. Frecuentemente, la formulación más natural será decir que siento placer porque, por ejemplo (como supongo), este autor ha llegado, y no que siento agrado por la (supuesta) llegada de este autor. Por otro lado, puede ser apropiado expresarlo como actividad: podría estar recibiendo placer al mirar el supuesto Giorgione más que en el supuesto Giorgione. En estas diversas formulaciones surgen

algunas diferencias genuinas, y el tipo de distinción que he estado trazando necesita refinamientos para tratar con ellas, pero la expresión del placer a partir de sus objetos, que es lo que estoy discutiendo aquí, no siempre se reduce al lenguaje del “porque”, y generalmente es una posible alternativa a él, y, como he tratado de argumentar al principio de este artículo, la forma “porque” es una causa de perplejidad filosófica que la otra formulación, propiamente entendida, puede ayudar a resolver. La expresión del placer como actividad de nuevo parece más independiente, y en algunos casos no se reduce a su formulación a partir de los objetos, a menos que en aquellos casos la actividad sea en sí misma el objeto de placer en el mismo sentido. Sin embargo, como se argumentó en la discusión del caso de la herencia, es también legítimo que la expresión del placer a partir de su objeto o del “porque” no pueda ser reducida en todo caso a su expresión como actividad. Ya no discutiré más aquí su formulación como actividad, que sospecho está más estrechamente relacionada con el concepto de disfrute que con el de placer. En cualquier caso, lo que nos ocupa ahora son los problemas de la creencia y el conocimiento en relación con el placer, y éstos surgen mucho menos con las actividades. En la medida en que esto es así, puede ser que algunas de las presentes explicaciones puedan adaptarse para tratar con ellas. Así que parece genuino decir que si estoy dedicado a una actividad o llevo a cabo una acción, hay un sentido en el que debo saber lo que estoy haciendo, pero esto sólo significa que hay alguna descripción de lo que estoy haciendo bajo la cual sé que estoy haciéndolo, y puede haber muchas otras que podría ofrecer de lo que estoy haciendo, que de hecho no se aplican. Podría estar bajo una de estas últimas descripciones en que estaba disfrutando hacer lo que estaba haciendo (caso *c*), o bien la descripción errónea podría no importar porque no era *como esto*, que la acción o actividad se estaba disfrutando (caso *b*), etcétera.

La analogía, por un lado, entre los casos discutidos de señalamiento y atención, y, por el otro, los del placer, no se detiene aquí. Analogías similares se pueden encontrar para los casos *e* y *f*:

- e*) Puedo sentir placer por alguna característica que sé que es ilusoria, pero que, no obstante, me agrada. Así, debido a mi miopía, puedo encontrar una pintura estridente como agradablemente amortiguada; si tengo una fiebre alta, puedo encontrar una bebida ácida placenteramente dulce; después de la mezcalina, puedo encontrar en mis viejas cortinas una excitante profusión de color.
- f*) Como *x* puede atraer mi atención y, no obstante, no saber qué es lo que hay en *x* que atrae mi atención, puedo sentir agrado por *x* y, no obstante, no saber qué tiene *x* de agradable. Nuevamente con el placer, como con la atención, uno encuentra en tales casos el fenómeno enigmático del ser aparentemente capaz de descubrir por reflexión lo que le agradó o llamó su atención.

Así que parece haber una analogía entre los casos de error, ignorancia e ilusión que pueden surgir con los objetos que atraen mi atención, y los que me producen placer. Sin embargo, antes de dejar la consideración detallada de la analogía, hay una característica

de ella que debe investigarse. Un modo notable en el que los errores pueden afectar los placeres es que mi placer puede estar fundado en creencias falsas acerca del pasado o el futuro. Así, pude sentir agrado porque supuse que la observación que había escuchado recientemente sin querer había sido un cumplido acerca de mí, aunque no lo era, o pude sentir agrado porque pensé que iba a ver a cierta persona al día siguiente, y no fue así. El caso mencionado de la supuesta herencia, es un tipo de amalgama de las dos, puesto que implica tanto la creencia de que alguien ha hecho un testamento a mi favor, como la creencia de que tendré el dinero; aunque es presumiblemente la segunda creencia la que está más estrechamente conectada con el placer. Ahora bien, con la herencia, como hemos visto, puede usarse, desde luego, la expresión del placer a partir de su objeto y este caso fue asignado a una cierta clase en la analogía, la clase *d*, de objetos que pienso que existen pero que son imaginarios.

Pero la situación es más complicada de lo que esto sugiere. En primer lugar, hay cierta asimetría entre el pasado y el futuro en este aspecto. En cuanto al futuro, debemos distinguir entre los placeres de anticipación y los placeres del evento. Yo puedo, por supuesto, tener los primeros sin que los segundos ocurran en absoluto, o bien porque el evento no sucede, o porque, aunque suceda, resulte no ser placentero. Estas dos últimas posibilidades vienen a ser lo mismo, por lo que concierne a la “carencia de base” de los placeres de anticipación, pues los placeres de anticipación consisten en la anticipación del placer.[34]

En cuanto al pasado, por consiguiente, es generalmente válido que el placer de la memoria consista en la memoria del placer. El “placer de la memoria” aquí no ha de ser confundido, desde luego, con el placer de la reminiscencia o, incluso, del recordar; con éstos, el placer se torna una “actividad” presente, a saber, recordar o tener reminiscencias acerca de algunos eventos pasados, que no hayan sido por fuerza particularmente placenteros. Estamos interesados en el “placer de la memoria” en un sentido diferente, el de la continuación o reavivamiento del placer asociado con algún evento pasado; pero incluso en esta conexión debemos distinguir dos modos muy diferentes en los que pueden surgir errores. El primero es cuando el placer es continuado o revivido por el recuerdo correcto de un placer originalmente basado en un error, como en el caso del supuesto cumplido. Esto no presenta ninguna dificultad especial: es simplemente, como quien dice, la herencia de un error. Pero hay otro caso, en el que, debido a un error de memoria, o bien “recuerdo” un evento placentero que no ocurrió y que no creo que hubiera ocurrido al supuesto momento del acontecimiento, o bien recuerdo mal; recuerdo como placentero algo que sucedió pero que para mí, en su tiempo, no fue placentero; en cualquier caso, puedo sentir un placer presente ante el supuesto evento pasado. Ésta no es la herencia de un error, sino un error de herencia. De éstos, el segundo es más similar al caso de un error acerca del futuro.

Ahora bien, no es del todo claro cómo ha de tomarse la formulación del placer a partir de los objetos en tales casos, ni cuál será su analogía con la atención. Los casos pueden agruparse, como se sugirió, en la clase *d*, con lo que formarían parte del grupo en el que mi atención es atraída por lo realmente inexistente, pero esto no es muy

iluminador. No obstante, al extender un poco la noción de atención, se pueden encontrar algunas analogías mejores. Sólo continuar o revivir un placer que surgió originalmente de un error, es como tener mi atención atraída por algo y seguir concentrándola en ello cuando el objeto mismo, o más bien lo que supuse que era este objeto, ya no existe. Más difícil es el caso del placer presente basado en la mala memoria; no obstante, aquí se puede hallar una analogía con alguien que dirige su atención a algún supuesto suceso, o éste, o quizá su memoria, que se impone a su atención. Con el futuro, mi atención puede, desde luego, ser dirigida a un supuesto evento futuro, como lo hago en mi expectativa o espera, que (como he observado arriba) puede ser previo a poner mi atención en la cosa real una vez que aparezca. Esto, aunque la analogía no es perfecta, es algo como el placer de anticipación seguido por el placer del evento. El punto requiere ulterior investigación, pero me gustaría sugerir que la distinción entre los objetos de atención “invitados” y “no invitados” tienen otras aplicaciones en conexión con el concepto de placer. Por ejemplo, está conectada con la distinción, trazada primero por Platón,[35] entre los placeres que consisten en la satisfacción de un deseo, y los que no.

En el curso de esta discusión he trazado algunas analogías entre, por un lado, la relación del placer con sus objetos y, por el otro, la de la atención con los suyos. Se puede sugerir ahora que estas analogías no son meramente analogías: existen porque la atención está relacionada con el placer, y debido a que la relación del placer con sus objetos *es* la relación de la atención con sus objetos. Si me agrada algo, eso atrae mi atención, y mientras más me agrada, más atraerá mi atención. Se puede subrayar aquí que hay quizá casos de placer que no tienen objeto; en los que uno simplemente se siente pleno de bienestar. De hecho, la palabra “placer” parece muy raramente usada en tales casos; se emplea más bien “alegría”, “satisfacción”, etc.; o si lo es, al “placer” se le da un pseudo-objeto (“el placer de vivir”). Pero si son casos de placer, es interesante que también son casos en los que habitualmente mi atención no está dirigida a nada en particular.

Esto no significa, desde luego, que recibir placer sea simplemente fijar la atención en algo; que encontrar algo placentero sea sólo que atraiga la atención y se sostenga; que los placeres de anticipación sean simplemente anticipación; que la atención pueda, asimismo, simplemente dirigirse a, o sostenerse por lo no placentero. Más bien el placer es un modo o una especie de atención.

Esta conexión entre placer y atención ha sido notada antes, *e.g.*, por el profesor Ryle. [36] Pero Ryle no discute cuál es la relación de la atención con sus objetos, en particular con objetos de los cuales erróneamente se cree que existen. He tratado de mostrar que el concepto de atención es en sí mismo lo suficientemente complejo en estos aspectos para iluminar las correspondientes complejidades del concepto de placer. Ahora bien, podría justamente preguntarse qué tanto explica esto, pues, puede decirse, puesto que no toda atención es atención de lo placentero, que todavía falta explicar la real relación del placer con la atención. Esto es verdadero, pero al menos la relación del placer con sus objetos puede ser algo clara si se muestra que esta relación es la de la atención y sus objetos, y si esta relación explica algo. Lo último que he tratado de hacer, a grandes rasgos, al sugerir

que la idea de darle atención a una cosa puede estar basada en, aunque no directamente derivada de, la simple noción de señalarla. En particular, espero que la introducción de la noción de que una cosa atrae mi atención sobre sí misma puede ayudar a explicar cómo, aunque el placer es algo que me sucede, más que algo que hago, a pesar de la caracterización de sus objetos, comparte rasgos epistemológicos con la caracterización de los objetos señalados, que es algo que hago, más que algo que me sucede. Estas categorías de “algo que hago” y “algo que me sucede” son, desde luego, demasiado burdas, pero es significativo que sean demasiado burdas para los conceptos tanto de placer como de atención, y por mucho en los mismos modos. Por ejemplo, a veces ocurre que algo atrae mi atención inesperadamente porque, de hecho, he estado inconscientemente mirándolo, y algo muy parecido puede ocurrir con el placer.

El tipo de relación que se da con un objeto de la atención o del placer, que he estado tratando de investigar, es uno de los que algunos filósofos han investigado bajo el título de “intencionalidad”, y aunque deliberadamente he evitado la palabra, espero que las observaciones presentes puedan sugerir una línea para aclarar esta oscura noción. Algunos de estos filósofos (quizá Husserl) al parecer han sostenido que cada tipo de “estado de conciencia” era al final inanalizable. Si pudiera aclararse no sólo la relación entre la atención y sus objetos, sino la del placer con la atención, podría encontrarse que esta perspectiva es demasiado pesimista.

IV. Conocimiento y razones

Uno de los propósitos de este artículo es hacer algunas sugerencias acerca del papel que desempeñan las razones en el conocimiento. El otro es bosquejar un enfoque del tipo de conocimiento que pondrá esta cuestión en una perspectiva correcta. Este esbozo será, desde luego, incompleto, y la mayor parte de lo que habré de decir, esquemático. Mi intención es poner los factores principales en lo que me parece es la forma correcta en general.

1. Conocimiento proposicional y práctico

Me ocuparé solamente de lo que llamaré conocimiento *proposicional*, conocimiento que en el uso habitual del lenguaje es la confiada aseveración de verdades, y donde la pretensión de que es conocimiento lo que se está expresando implica como condición necesaria que lo que se afirma sea verdadero. Esto contrasta con el conocimiento *práctico*, cuya expresión paradigmática es la ejecución exitosa y habilidosa de alguna tarea. (Sin embargo, el éxito no está relacionado con el conocimiento práctico como la verdad lo está con el conocimiento proposicional: sin la verdad de la proposición no hay conocimiento proposicional, pero el conocimiento práctico puede estar presente y activo, aunque por causas extrínsecas no tenga éxito.)

No es fácil formular la distinción y ésta admite casos indeterminados; sin embargo, es a la vez genuina e ineliminable, ya que ninguno de los dos tipos de conocimiento se reduce al otro; supondré esto aquí sin argumentarlo. Haré dos observaciones adicionales acerca de la distinción, ambas relevantes para lo que sigue. Primeramente, uso el término *proposicional*, y no el generalmente aceptado, *teórico*, porque en el segundo suele incluirse la noción de lo sistemático. Algún conocimiento proposicional es, desde luego, teórico, en el sentido de ser general, sistemáticamente ordenado, tener una estructura de leyes, etc., pero mucho del conocimiento proposicional no lo es. La idea de que lo que contrasta con la práctica es, en *ese* sentido, la teoría, puede generar una seria confusión (presente, pienso, en las reflexiones de Michael Oakeshott en su *Rationalism in Politics*, y en otras obras). En segundo lugar, no pienso que la distinción sea muy felizmente etiquetada, como lo hiciera Ryle, como una distinción entre *saber qué* y *saber cómo*. Estas etiquetas son, desde luego, correctas si son simplemente eso, pero su uso puede ocasionar que se pasen por alto varios hechos, algunos de ellos importantes. Por ejemplo,

no toda adscripción de conocimiento proposicional necesita, o incluso puede, adoptar tal como es la forma “*A* sabe que...”. “Sabe” puede emplearse en gran medida en cuestiones indirectas (como, en particular, en, “*A* sabe si...”). La relevancia filosófica de este hecho aparecerá en el siguiente apartado. Por otro lado, “*A* sabe cómo...” puede representar conocimiento proposicional (*A* sabe cómo funciona una estación de energía nuclear”); lo que representa el conocimiento práctico es, más bien, “*A* sabe cómo...” Pero entonces “cómo” no representa ninguna peculiaridad: el conocimiento práctico puede ser igualmente representado por “*A* sabe cuándo..., qué...” etc. Por consiguiente, algunos casos de “*A* sabe que...” pueden representar conocimiento práctico, si éstos después incluyen “para” (o una construcción equivalente) y (algo como) un demostrativo: “*A* sabe que este es el momento de agregar la sal”, “*A* sabe que el que hay que usar es el que tiene este aspecto” etcétera.

Posiblemente debería añadir que no considero que el idioma inglés defina con exactitud el concepto de conocimiento; pero lo anterior puede servir para recordarle a uno las posibilidades equívocas de las mismas etiquetas extraídas de ese lenguaje.

2. “Saber qué” y la situación del examinador

Los filósofos que se han concentrado en la tercera persona (y no, como lo hiciera Descartes, en la cuestión de “¿qué sé yo?”, un enfoque que plantea sus propios problemas) han tendido a reducirse a la cuestión de las verdades de hecho como “*A* sabe que *p*”. Es decir, han buscado condiciones suficientes para dar una respuesta afirmativa a la pregunta “¿sabe *A* que *p*?” En la vida real, una implicación muy natural al plantear esa pregunta es que quien la hace sabe que *p*, y está preguntando si *A* lo sabe. Al tomar a esa pregunta como central, un efecto inmediato, que tiene esa implicación, es restarle importancia a cierto tipo destacado de casos, a saber, aquellos en los que quien pregunta no sabe cómo se hallan las cosas con respecto a *p*, y quiere encontrar a alguien que sí lo sepa; una situación en la que podría preguntar “¿quién sabe si *p*?”, o “¿sabe *A* si *p*?”. Tenemos ya un enfoque que nos ayuda a olvidar un hecho banal pero importante: los conocedores son, para otros, fuentes de información.

Sin embargo, incluso al aferrarse a “*A* sabe que *p*”, el enfoque tiende a concentrarse solamente en una de varias direcciones. La favorecida es aquella en la que el objeto de la pregunta es saber si la relación cognitiva de *A* con la verdad de *p* es adecuada; en particular, el caso en el que se admite que *A* está convencido de la verdad de *p*, y la cuestión es si esa convicción está bien fundamentada. Esta situación, y más generalmente la situación en la que interrogadores informados se preocupan por las credenciales de *A* con respecto a una pieza de conocimiento, podríamos llamarla la *situación del examinador*. Está lejos de ser típica en la práctica. Por ejemplo, otra situación frecuente con respecto a la pregunta “¿sabe *A* que *p*?” es aquella en la que el objetivo es preguntar si la información de *p* ha *llegado* a *A*, o se ha *aproximado*: uno de nuestros intereses en el conocimiento es adónde ha llegado. Cuando tal es nuestro interés, el énfasis en las

credenciales (cuya relevancia es autoexplicativa para la situación del examinador) es menos importante. Cuando mi interés es si A ha llegado a saber algo que, por ejemplo, sería mejor para mí si no lo supiera, la cuestión central es si él ha recibido la información y la cree, y sería menos importante si las razones con las cuales A apoyaría su creencia son fuertes o adecuadas. No obstante, incluso en tal caso, hay un contraste en A simplemente por haber llegado a creer la proposición en cuestión, un contraste incrustado en el pensamiento de que le ha “llegado” la información. Si ha llegado a saber, por lo menos se trata de que simplemente no ha conjeturado; debe haber una ruta por la cual la información lo ha alcanzado, y el hecho de que lo que él cree es verdadero debe haber contribuido de algún modo a que haya llegado a creerlo. En esta última consideración, no importa cuán oscura sea, veremos una condición central del concepto de conocimiento.

Regresemos a los casos en que nuestro interés en lo que otra persona sabe tiene que ver con un asunto acerca del cual no tenemos conocimiento: nuestra preocupación (a diferencia de la del examinador) es averiguar, no tanto acerca de esta persona, sino del asunto disponible. Si p representa alguna proposición que puede ser utilizada para hacer una proposición de hecho, tomemos $pq-p$ para representar en general preguntas directas e indirectas (con las modificaciones gramaticales necesarias) que pueden ser formadas a partir de esa proposición. El ejemplo más simple de una cuestión indirecta, disponible en todo caso, es “si p ”. Pero hay, desde luego, muchas otras y varias posibilidades; si p menciona algún tiempo (“el tren sale a las 15h”) $pq-p$ puede ser una pregunta de “cuándo” (“¿cuándo sale el tren?” “... cuando el tren salga”); si p menciona un lugar, $pq-p$ puede ser una pregunta de “dónde”, etc. Me parece que todas las siguientes proposiciones son verdaderas:

- a) En muchas situaciones comunes, todo lo que es necesario para que se de el caso en que A sabe $pq-p$ —además de estar en realidad en lo correcto en este caso (véase b más adelante)— son cosas como que A se halla casi siempre en lo correcto acerca de asuntos de este tipo, porque, *e.g.*, en el asunto del tren, él se ha aprendido el horario, o que sabemos, lo que quizá A mismo no sabe, que A ha llegado a sus creer esto por medios confiables. Así que puede ser que A haya creído porque B le dijo, y sabemos que B es una autoridad confiable en estos asuntos, aunque el mismo A puede no serlo, sin haber reflexionado (por ejemplo) sobre esa cuestión.
- b) Si las condiciones de este tipo se satisfacen, entonces todo lo que es adicionalmente necesario para que sea el caso que A sabe $pq-p$ es que las creencias que tiene con respecto a esto sean verdaderas, *i.e.* que, de hecho, se halle en lo correcto. (Supondré a lo largo de este artículo que la *creencia*, al implicar un considerable grado de convicción, es una condición necesaria del conocimiento. Dudo que esto sea verídico sin condiciones, pero la suposición servirá para el presente análisis.)
- c) Si A sabe $pq-p$, y P es la clase de respuestas correctas a la pregunta $¿pq-p?$, entonces para algún miembro de P , q , A sabe que q . Esta formulación más bien incómoda pretende dar lugar a hechos tales como que si $¿pq-p?$ es una pregunta —“quién”, digamos de la forma “¿quién hizo X ?”—, A puede saber de cierta persona, por

algunas descripciones, que él hizo X , pero no a partir de otras descripciones.

Si las proposiciones $a-c$ son todas verdaderas, entonces es posible que A sepa que q sin ser el caso que A pueda repetir razones, o al menos razones adecuadas, para q . Pues los tipos de condiciones mencionados en a no implican necesariamente nada, o nada muy sustancial, acerca de la *conciencia* que A tiene de las razones para q , o de su propia relación con la verdad de q . Son condiciones *acerca de A*, más que condiciones *sobre A*; podemos llamarlas (para usar una frase que debo a una discusión con Mr. G. O'Hare) condiciones *externas*. Si $a-c$ son verdaderas, entonces la satisfacción de tales condiciones externas puede, junto con la creencia verdadera, ser suficiente para el conocimiento.

3. Conocimiento sin razones

Uno podría esperar que esto fuera así, pues si consideramos el clásico análisis tripartito del conocimiento proposicional, a saber, que A sabe que p si y sólo si *i) p* es verdadero, *ii) A* cree que p , *iii) A* tiene buenas razones para p , hay una dificultad notoria, que es probable que se genere un retroceso, y en más de un modo.[37] Pues si la razón de A para p está constituida por alguna otra proposición que apoya p , entonces parece que ésta a su vez debe ser algo que A sabe, a menos que lleguemos a la no muy atractiva conclusión de que la diferencia entre creencia verdadera y conocimiento es básicamente la diferencia entre creer una cosa verdadera y creer dos cosas verdaderas que están conectadas. Más aún, el hecho de que esta otra proposición apoye p debe en sí mismo (parece) ser algo que A sabe, lo cual genera un retroceso adicional.

Algunas formulaciones de la tercera condición (como la de Ayer: “ A tiene el derecho a estar seguro”, en *The Problem of Knowledge*) están diseñadas para ser lo suficientemente generales como para tomarse en cuenta, al menos en algunos casos en que la fundamentación apropiada de la creencia verdadera de A no asume la forma de otra proposición que A sabe o cree. Pero el propósito de tales intentos ha sido generalmente acomodar alguna clase o clases de proposiciones que supuestamente poseen algún carácter especialmente evidente o básico: los *fundamentos* han de frenar el retroceso. Cualquier cosa que se diga acerca de los fundamentos del conocimiento, no vale la pena traerla a colación, como una respuesta a esta dificultad en relación con la explicación de lo que el conocimiento es en general. Más bien, debiéramos reconocer que si estamos hablando *en general* acerca del conocimiento, no solamente no es necesario que el conocedor sea capaz de apoyar o fundamentar su creencia verdadera refiriéndose a otras proposiciones, sino que no es necesario que se halle en ningún estado especial con respecto a esta creencia en absoluto, al menos cuando se trata de lo que él puede conscientemente repetir. Lo que es necesario —y lo que representa el hecho indudable de que el conocimiento difiere de la mera creencia verdadera— es que debe satisfacerse una o más de una clase de condiciones, las cuales relacionan el hecho de que A tiene esta creencia tomando como un hecho que la creencia es verdadera: condiciones que se

pueden resumir de la mejor manera por la fórmula de que, dada la verdad de p , no es un accidente que A crea que p en vez de $\text{no-}p$. Esta fórmula es vaga y generosa en exceso pero nos conduce, pienso, en la dirección correcta; en particular si consideramos que la noción de ser “no accidente” es básicamente la misma noción que se emplea en una investigación causal.

Supóngase que se piensa que una médium espiritista o algo por el estilo tiene poderes de clarividencia. Hace aseveraciones, de las cuales está convencida, acerca del paradero de ciertos objetos, o de determinadas situaciones en algún otro lado. Supongamos que sus aseveraciones son claras y determinadas y, más aún, verdaderas. Se pueden plantear ahora dos preguntas: ¿es simplemente un accidente, extraordinaria buena suerte de su parte, que se halle en lo correcto en esta serie de casos?, y ¿sabe ella? Mi tesis es que éstas son la *misma* pregunta. Si estamos convencidos de que no puede ser sólo suerte, que la probabilidad de un éxito casual a este nivel es demasiado pequeño, entonces surge la pregunta de cómo se logró el éxito, que es lo mismo que preguntar ¿cómo lo sabe? (La diferencia en inglés, que preguntamos *why he believes* [por qué cree], en lugar de *how does he know* [cómo lo sabe], subrayada por Austin y otros filósofos, es, desde luego, indicativa de una verdad). Quizá no seamos capaces de descubrir cómo sucede esto; en ese caso, no podremos decir cómo lo sabe. Si se nos deja en esa posición, no podremos (*pro tanto*) negar que ella sabe; la correlación, desde luego, existe, pues es demasiado buena para que sea casualidad, y exige una explicación. Simplemente se nos dejaría admitidamente perplejos con el hecho de que ella sabe, pero hay una completa confusión acerca de cómo es que sabe.

Podemos tomar también un conocido ejemplo de la vida real. Alguien estrechamente relacionado con otro por lo regular sabe cómo se siente éste, qué pensamiento le ha llegado, cómo reaccionará, por qué reaccionó de cierto modo. Sus bases para estas convicciones son con frecuencia, en el nivel consciente, virtualmente inexistentes o al menos irremediamente inespecíficas, pero si es verdad que tal persona suele estar en lo correcto, y no, como ocurre a veces, sólo revisando los ejemplos negativos, entonces vacilaremos poco para atribuirle conocimiento. En este caso también tenemos una idea general de cómo es que saben: hay un esquema explicativo disponible, el cual podría en principio cumplir, sin duda, mediante una investigación cuidadosa de qué es lo que les está sucediendo subconscientemente.

Vale la pena decir, de paso, que existe tal conocimiento no razonado; más aún (aunque éste no es el punto que nos ocupa), de conocimiento personal y social que el sujeto no puede ni siquiera expresar adecuadamente, menos aún justificar. Esto no se opone, de ningún modo místico, a una imagen racional y científica del mundo. Que los humanos puedan entender lo humano de maneras en que lo no humano no pueda ser entendido, no muestra que el hombre se encuentre fuera de la naturaleza; más bien revela el tipo de sistema en la naturaleza que es él. Pensar, más aún, que, a menos que el entendimiento humano no razonado sea mágico, debe ser posible *reemplazarlo* con procedimientos articulados y razonados de forma científica, es nuevamente un *non-sequitur*, y en una buena parte una ilusión científicista. Insistir en pensar acerca de

situaciones personales en tales términos hace, después de todo, una diferencia psicológica concreta en el pensador; una diferencia que puede perfectamente bien destruir, sin sustituir, el conocimiento que de otra manera tendría. El papel e importancia de lo reflexivo y autoconsciente en los asuntos humanos es, desde luego, algo serio, pero no lo es pensarlo como si se estuviera remplazando la adivinación del clima con la meteorología.

4. La dificultad de producir un criterio

Sostengo, entonces, que no es un requerimiento *en general* del conocimiento que el conocedor tenga que estar en algún estado consciente especial en relación con su creencia verdadera; el requisito es más bien que no sea accidente, concedida la verdad de p , que él crea p más que $\neg p$. Pero si esto es considerado como un *análisis* del conocimiento, no hay duda de que no servirá, puesto que es demasiado vago y (en interpretaciones naturales) demasiado generoso. Pues supóngase que A , siendo de Guinea, dice a B falsamente que es de Ghana; pero (supongamos de una manera extravagante), debido a que el inglés hablado por A tiene acento guineano, B entiende “Guinea” cuando dijo “Ghana”. Entonces B ha creído verdaderamente que A es de Guinea, y (en un sentido obvio) no es accidente, relativo a que A es de Guinea, que esto haya ocurrido, pero difícilmente se podría decir que B ha adquirido conocimiento de este modo, en contraste (por ejemplo) con una situación en la que, familiarizado con el acento guineano, ve a través de esa falsificación. O, nuevamente, el contador de una empresa, deprimido por razones personales, da por ello una imagen deprimente de los asuntos de la empresa. El gerente se deprime por ello; cuando está deprimido, tiende a pensar que todos los demás también lo están, y forma la creencia (verdadera) de que el contador lo está. También esto puede difícilmente contar como la adquisición de conocimiento.

Éstos y otros casos parecen ser contraejemplos de la explicación de “no accidente” adoptadas sin mayores restricciones como un análisis, puesto que parecen satisfacer (a menos que secretamente supongamos alguna restricción no explicada) la condición general de que la adopción de la creencia no era accidente, dada su veracidad y, no obstante, éstos no son casos de conocimiento. Se puede esperar que una ulterior investigación provea algunos tipos apropiados de restricción, tales como aquéllas necesarias para distinguir el paso de una pieza de información (en un sentido proposicional, no en el sentido generalizado usado, *e.g.*, en las ciencias biológicas, aunque esto no es irrelevante) de la más general noción de una cadena causal con la misma proposición en cada fin de la misma. En ausencia de las requeridas restricciones, ofrezco la cláusula del “no accidente” no como parte de un análisis sino, como dije antes, como una etiqueta para cierto tipo de condiciones, cuyos requerimientos generales necesitan ser explicados con mayor precisión.

Un requisito general, desde luego, es que la ruta que va del hecho a la creencia debería *en general* ser productora o preservadora de la verdad; que las creencias

engendradas por este *tipo* de proceso deberían tener una buena probabilidad de ser verdaderas. Pero eso, si bien es correcto, en realidad no nos lleva muy lejos, pues la pregunta más difícil permanece: ¿qué nivel de generalidad de descripción ha de tener un proceso o tipo de proceso para poder decir que tiene o no la característica productora de verdad? O, formulado de otra manera, ¿qué ha de contar como el mismo, o diferente proceso? Así que, puede decirse con respecto a los dos casos que he mencionado, que lo que está equivocado es que los procesos de producción de creencia implicados —oír mal en un caso, estar influido por un estado de ánimo en el otro— no son, en general, procesos productores de verdad. Bien, considerado en este nivel de generalidad, no lo son, pero considerados en un nivel más específico, con las características distintivas de los casos incluidas, sí lo son, y la pregunta es por qué el primero y no el segundo nivel de descripción debería representar propiamente nuestra visión de la situación.

Problemas muy similares surgen con la caracterización de las *razones*. Un hecho que, descrito en un cierto nivel, constituye una buena razón para creer algo (e.g. “dicho por Jones, una autoridad confiable”) puede tener muchas descripciones bajo las cuales aparecerá como una mala razón, o como ninguna razón en absoluto (“alguien me dijo”, “me dijo un hombre que nunca ha estado en París”). Este nada sorprendente paralelismo entre los problemas de encontrar la correcta aplicación del “no accidente”, y los problemas de evaluar las razones, es uno de los varios fundamentos para suponer que el tipo de dificultad presentada por estos casos no concuerda con la explicación de “no accidente”, sino que solamente exige un refinamiento.

5. El lugar de las razones

Nada de esto es para negar que la posesión de razones desempeña un papel importante en la economía del concepto de conocimiento. Sería altamente paradójico si no tuviera tal papel. Pero esta explicación nos permite ver la naturaleza y enfoque de ese papel más claramente que una explicación que insiste simplemente en la posesión de razones.

En muchos casos será altamente probable, y en otros necesario, que el único modo en que un sujeto pueda con seguridad y (relativo a la verdad) no accidentalmente adquirir creencias verdaderas sea por medio del *pensamiento* de consideraciones que apoyen la pretensión de verdad, es decir, por medio de razones. Es probable que sea así con el conocimiento de asuntos espacial y temporalmente lejanos al sujeto, aunque aquí vale la pena recordar la perogrullada de que todos poseemos información acerca del pasado (y dejando de lado el caso especial de nuestros propios pasados), que es justamente aceptada como conocimiento, aunque las proposiciones que emanen de ella no puedan sustentarse con razones lógicamente adecuadas, sino en el hecho “externo” de que hemos adquirido la información a partir de fuentes presumiblemente confiables que ahora no solemos recordar. Un tipo especial de requisitos para las razones aparece en proposiciones de ciertos tipos lógicos. Concerniente a las proposiciones matemáticas, por ejemplo, con excepción de las del tipo más simple, está la fuerte presión detrás de la

concepción platónica de que la distinción entre conocimiento y creencia verdadera se halla en la posesión de un *aitias logismos*, una cadena de demostración. Platón mismo sostuvo (por lo menos en el *Menón*) que la razón de este requisito, y de allí la superioridad del conocimiento, reside en la mayor permanencia de creencias atadas de esa manera. Puede o no ser psicológicamente verdadero que tales creencias sean más permanentes, pero esto difícilmente va al corazón del asunto, que es seguramente más al punto de que el acceso a la verdad matemática debe necesariamente encontrarse a través de la demostración, y que por lo tanto la noción de creencia verdadera no accidental en matemáticas esencialmente implica la noción de demostración matemática (los puntos en los que el modelo platónico de recuerdo sirve precisamente para confundir).

Ahora bien, puedo realmente creer en una proposición matemática aunque no la pueda demostrar, porque se me ha dicho con autoridad que es verdadera. Habría un amplio consenso en que esto no constituye conocimiento: conocimiento, esto es, de que p , donde p es la proposición matemática. Pero hay otra pieza de conocimiento que bien podría decir que tengo en estas mismas circunstancias: el conocimiento de que p es una verdad matemática. Parecería paradójico separar en cualquier modo saber que p , y saber que p es verdadero. Pero si estas dos se mantienen lógicamente vinculadas, entonces debe aparecer una línea entre saber que p es verdadero, y saber que p es una verdad de una ciencia dada. Ser verdad de una ciencia dada es ser, en un sentido especial, parte del *conocimiento*, y saber que una proposición tiene esa condición es saber que ha de ser contada por los principios de esa ciencia como *conocida*, y no que sea, por ejemplo, sólo una conjetura plausible.

Esta consideración nos trae a un tipo final y diferente de conexión entre conocimiento y razones. Hasta aquí he estado discutiendo el conocimiento como, en un cierto sentido y en parte, un concepto psicológico que ha de aplicarse a individuos con respecto a su asirse a proposiciones verdaderas, pero existe algo conocido como conocimiento impersonal, como cuando hablamos del estado del conocimiento en un campo dado, o de algo que se sabe ahora que no se sabía hace 20 años, o que no se conoce la estructura de cierta molécula aunque existan teorías o hipótesis acerca de ella.

Que debería haber un conocimiento absolutamente impersonal parece, a primera vista, imposible: si p es conocido, entonces alguien debe seguramente conocerlo, pero esta aparente perogrullada se halla en conflicto con otras cosas que también estamos dispuestos a aceptar, y al menos una de ellas se tiene que abandonar. Permitamos que $k_i p$ represente “se sabe que p ”, en el sentido que hemos analizado, y $k_a p$ represente “ a sabe que p ”. La perogrullada aparente es:

1) Si $k_i p$, entonces, para algún a , $k_a p$.

Lo siguiente, sin embargo, parece ser plausible:

2) Si $k_i p$ y $k_i q$, entonces $k_i(p \text{ y } q)$.

Pero de hecho esto que sigue no es verdad:

3) Si kip y kiq , entonces, para algún a , $ka(p$ y $q)$;

Si 3 fuese verdadero, la era del sabio universal no se habría quedado detrás, pero 3 está implicado por 1 y 2, puesto que “ p y q ” puede, por supuesto, ser un caso de sustitución de p en 1. No exploraré este problema aquí; sólo quiero señalar que tenemos que resistirnos a la tentación de defender 1 basándonos en las dos suposiciones: *a*) que la lógica de kip se revelará concentrándose sólo en aserciones verdaderas de la forma kip , *b*) que cualquiera que verdaderamente afirme o, al menos, que sepa, que kip debe por sí mismo saber que p . Ninguna es verdadera. Una comprensión adecuada de kip tomaría en cuenta, por ejemplo, preguntas de si kip ; y por lo que a b concierne, ya hemos visto en el caso matemático que es falso.

He mencionado este problema solamente para mostrar que las relaciones entre el conocimiento impersonal y el personal no son tan simples como podría parecer a primera vista. Ahora bien, el conocimiento impersonal de hecgi tiene un compromiso especial con las razones. Los conjuntos de conocimientos son esencialmente, a mayor o menor grado según la materia, sistemáticos. Hay una razón pura y otra aplicada para esto. Pura, porque la intención no es sólo saber sino entender, y, al menos en el caso de la ciencia, entender necesariamente implica organización y economía. Aplicada, porque un conjunto de conocimientos será sólo libremente asequible y abierto a crítica si está racionalmente organizado, y hay, por supuesto, otras consideraciones que apoyan el mismo punto. Por lo tanto, el conocimiento, en este sentido, debe tener razones; puede incluso tener fundamentos, aunque esto es una pregunta abierta. Sin embargo, lo que es una pregunta abierta *de*, pertenece al campo de la filosofía de la ciencia; estos son asuntos diferentes de los de la epistemología en general, es decir, del estudio del conocimiento como tal, y particularmente del conocimiento personal.

Dígase lo que se diga acerca de las relaciones entre conocimiento personal e impersonal, es por supuesto verdad que lo que se sabe es fragmentariamente sabido por varias personas; el *savant* ha interiorizado alguna parte de un conjunto de conocimientos. Siendo esto verdad, su conocimiento personal satisfará los estándares de organización racional que son apropiados a un conjunto de conocimientos, pero no debemos tomar ese caso especial *como la* clave de la explicación que deberíamos dar sobre el conocimiento personal, tal cual.

V. Identidad e identidades

En 1955 apareció una ingeniosa y agradable novela de Nigel Dennis llamada *Cards of Identity* [Tarjetas de identidad]. En ella habla de una organización llamada “Club de Identidad”, la cual se dedica a rehacer a las personas dándoles un nuevo pasado y un nuevo carácter; una nueva identidad. Había mucha discusión acerca del *nombre* que cualquier personaje dado debería de tener. Lo siguiente da una idea:

—¿Ha estado con nosotros durante mucho tiempo?

—Muchos y largos años. Vino de la armada directamente aquí. Lo encontré, muerto de borracho, en un barrio bajo de Portsmouth.

—Y ¿le gusta su nombre?

—Se apegó a él inmediatamente. ¿Me puedes explicar?

—Me gustaría. Empezamos con la premisa de que todo mayordomo cree que ha nacido para comandar una flota.

—Eso es correcto. Prosigue.

—Nelson, sentía uno, era un nombre demasiado común... Pero en Jellicoe encontrabas *todo* —una sugerencia belicosa, resonante, desafiante, discretamente balanceada por una sílaba de apertura indicativa de una naturaleza congelada e insegura. En pocas palabras, aunque él será para siempre parcialmente algo rosa, y se la pasará temblando de culpa sobre un plato, tiene, en su totalidad, madera para liderazgo.

—De primera, Beaufort. Y así era, exactamente. Incidentalmente, te puede interesar saber que primero jugué con la idea de una identidad tomada de la pista de carreras. Pero cuando hice unos cuantos sondeos gratuitos, se encogió de horror. Es importante saber eso, a propósito. Nunca, a excepción de algunos casos raros, hay que construir sobre el disfraz existente. Imagine el horror de este desdichado hombre si hubiera tomado su apariencia con todo y pantalones de pana y lo hubiera llamado Donoghue.

—“*Demasiado* irlandés”, murmuró la señora Mallet adormilada. “No la broma que queremos aquí de momento, con tanto que hacer”. [38]

En relación con esto, “darle una nueva identidad” significa convertirlo en cierto tipo de persona. Aunque no es nada más un mayordomo dominante, sino un mayordomo dominante llamado Jellicoe, nacido en cierto pueblo cierto día, etc., sigue siendo un tipo de persona, aunque muy individualizado. Subyacente a dicho modelo, en el sentido de la identidad que le es aplicada a cierto ser humano, tenemos la idea de ese ser humano y su identidad particular. La persona que recibe la nueva identidad tipificada, todavía retiene su antigua identidad en el sentido particular: él es, exactamente, el mismo ser humano que fue encontrado en el barrio bajo y fue transformado en un orgulloso mayordomo por el Club de Identidad.

Muchos problemas filosóficos acerca de la identidad tienen que ver con los criterios para identificar cosas particulares. En la primera parte de este ensayo consideraré diversos problemas filosóficos que están asociados con el concepto de la identidad de una

cosa particular. Esto conduce a un caso especialmente importante de la relación entre clases y ejemplares, que abordaré en la última parte del ensayo: ésta es la noción de identidad social de una persona.

Empiezo con cuestiones más estrictamente metafísicas. Identidad implica estrechamente contar, ya sea en forma sincrónica o al paso del tiempo, y los problemas de identidad están conectados con lo que, en términos antiguos, pueden ser llamadas cuestiones del uno y de los muchos, de cuántas cosas de cierto tipo hay en cierto lugar o por cierto periodo. Como Frege atinadamente insistiera, la pregunta “¿cuántos?” siempre exige una respuesta a “¿cuántos qué?” Lo que tenemos ante nosotros puede ser un bosque y 500 árboles, o una biblioteca y 2 000 libros. En estos casos, las cosas numerosas, ya sean árboles o libros, constituyen, hacen a, o son partes de una única cosa, y esta única cosa se ve como un agregado de cosas particulares. Pero a veces lo uno puede ser mucho más obvio que los muchos. Éste es un punto importante en biología. Hay muchas especies de animales, por ejemplo medusas de la orden *Siphonophora*, como la conocida aguamala, cuyo ejemplar individual visible es, de hecho, una colonia de individuos más pequeños y más bien de hecho diferentes individuos reunidos: hay una aguamala que es una colonia o asociación de muchos animales que la constituyen. Como los elementos de la colonia pueden estar funcionalmente diferenciados hasta cierto punto, esto plantea preguntas acerca de los límites de un individuo y de qué puede considerarse como una especie biológica. Con la aguamala el caso no es que una medusa forme parte de muchas medusas, y tampoco que forme parte de muchas aguamalas. En otros casos, sin embargo, el uno y los muchos pueden llevar el mismo nombre.

Esto es característico de la clonación. Hay una especie de arbusto del desierto que crece en un círculo creciente hasta que empiezan a morir partes en medio y partes entre las matas sobre la circunferencia. Este proceso puede dar lugar a un círculo, en algunos casos un círculo muy grande, de arbustos separados que son en cierto sentido un solo arbusto. ¿Es el resultado un arbusto, con partes espacialmente separadas, o varios arbustos? En este caso puede no importar mucho. Es fácil responder que es tanto uno como muchos, o que no importa lo que uno diga al respecto, pero uno no debería apegarse demasiado a la descripción de algo como “una cosa”. Se puede entrar en verdaderas dificultades, especialmente cuando estamos refiriéndonos a especies que se mueven. Las cuestiones de identidad, no nos sorprende, es común que se agudicen cuando las cosas se pueden mover.

Esto es obvio en el caso de la amiba. La amiba se divide y es extremadamente importante que su división constituya una forma de reproducción (asexual); el objeto de este proceso es producir dos amibas. Si hay dos amibas, entonces no es cada una idéntica a la amiba original. Si fueran idénticas a ella, serían idénticas entre sí, lo cual significa que serían solamente una amiba. Esto daría por resultado que, no importa cuán intensamente se reprodujera la amiba, su población nunca se incrementaría; simplemente habría partes de la amiba que aparecerían en diferentes lugares. Tenemos buenas razones para resistirnos a verlo de esta forma.

La división puede parecer como una manera particularmente buena de preservar la identidad de uno a lo largo del tiempo. Si pudiera dividirme, parecería que no solamente seguiría existiendo, sino que esto sería de manera doble. Pero, de hecho, la división destruye la identidad a través del tiempo. No es que la amiba aparezca en dos lugares a la vez, sino más bien que abandona su existencia para que haya dos amibas.

Un acertijo metafísico igualmente antiguo tiene que ver con la forma y la materia. Éste, de hecho, nos regresa muy rápido al uno y los muchos, y a un caso en el cual acabamos con más cosas de las que queríamos. El hecho de que una cosa de cierto tipo pueda estar hecha o formada por partes o piezas de materia, significa, lo que nos es suficientemente familiar, que las partes pueden cambiar mientras que el todo permanece siendo el mismo. En el caso de las criaturas, como nosotros, que se componen de células vivientes, que son casi todas remplazadas a lo largo de un ciclo periódico, algunos se ven tentados de decir “no somos realmente los mismos” al final de este periodo. Esto simplemente no es correcto, por una razón que muestra que el fenómeno en sí mismo no crea un acertijo. Un cuerpo vivo puede ser algo hecho de células, sin que sea el caso que el mismo cuerpo sea algo hecho de las mismas células. La conclusión es que el mismo cuerpo es el mismo algo hecho de células. Los acertijos no provienen realmente de esto, sino de casos que incluyen elementos peculiares, o nuevamente elementos ordinarios que tienen historias peculiares. Un paradigma de identidad a través del tiempo es provisto, y como pasa con frecuencia en la filosofía, por los objetos físicos comunes y corrientes, el tipo de cosa que J. L. Austin solía llamar “prendas de tamaño medio”. Hay muchos tipos de cosas que en comparación con aquellos objetos materiales son peculiares con respecto a la identidad. Tienen una historia y algún tipo de ubicación pero sus criterios de identidad parecen ser vagos o estipulativos: hay cosas como clubes, regimientos, ríos, etc. En estos casos es muy obvio que las partes que los constituyen cambian, pero no es claro lo que tienen de continuidad. Los regimientos se fusionan, los clubes cesan y reviven, y uno puede preguntarse qué significa que existan. Grupos como éstos pueden incluso ser pensados como algo con existencia discontinua, una idea que desde un punto de vista lógico es muy problemática. El río se seca un poco o, como los ríos en Australia, regularmente se seca durante periodos largos. Entonces ¿ya no hay río, o es un río seco?

No tiene sentido dar demasiada importancia a preguntas de este tipo. No quiero decir con eso, como los científicos y otros podrían estar dispuestos a pensar, que es filosofía pero no buen sentido; quiero decir que no es buen sentido filosófico. Pero no es siempre fácil apoyarse en la decisión convencional. Como vimos antes, con respecto al arbusto y la amiba, las cuestiones de identidad pueden pasar rápidamente de lo aparentemente trivial y verbal a lo que es genuinamente enigmático, y este hecho en sí mismo es revelador. Es sólo lo que se halla sobre la superficie de nuestro modo de hablar lo que puede ser fácilmente reacomodado.

Un ejemplo de esto nos lo da el famoso barco de Teseo, una nave en la cual se supone que el héroe regresó a Atenas y que fue conservada, aunque las piezas de madera con las que estaba hecha fueron remplazadas gradualmente hasta que no quedó una sola parte original. Esto dio lugar en los tiempos antiguos a una pregunta acerca de si el

producto final era el barco original. Thomas Hobbes brillantemente introdujo en este no muy interesante acertijo la idea de que si las piezas originales se hubieran conservado y que después se hubieran vuelto a armar, el resultado habrían sido dos candidatos a ser el barco de Teseo.[39] Como señalara Hobbes, y como podemos recordar del caso de la amiba, no podrían ser ambos ese barco particular. En el caso de la amiba, tenía dos descendientes, ambos de los cuales llegaron exactamente por los mismos medios; en el caso del barco, podría decirse que eran dos descendientes que habían llegado por diferentes medios, el de la forma y el de la materia.

Más que preguntarse cuál es el barco original, si es que alguno lo es, puede ser más interesante plantear la siguiente pregunta: ¿si cualquiera de estas cosas hubiera existido sin la otra, hubiese sido el barco de Teseo? Esto sugiere la idea de que la respuesta a tales preguntas de identidad ha de darse en términos del mejor candidato disponible. En estos términos, el barco continuamente reparado ciertamente será el original si el maderamen no ha sido mantenido y reensamblado, pero igualmente el barco hecho a partir del maderamen original, sería ese barco original si el otro hubiese sido destruido. Esto parece, a su manera, tener mucho sentido, pero hay algo contra-intuitivo acerca de la teoría del mejor candidato.[40] Esto conduce al resultado de que algo puede dejar de ser, o convertirse exactamente en lo mismo que un objeto original. El barco que Hobbes consideraba como el triunfador —el que estaba hecho con las tablas originales— era el peor candidato en tanto no era ni siquiera un barco, sino un montón de tablas. Cuando se reensamblaban para hacer un barco, ¿el otro dejó de ser el barco de Teseo? Más aún, si hay un concurso entre dos objetos que resulte en un empate, esto puede conducir a una elección arbitraria de uno de ellos para que sea la cosa original. Pienso que una teoría en estos términos es una explicación reconocible de algo, pero no de identidad numérica. Lo que más bien deberíamos decir acerca del ejemplo de Hobbes es que la descripción “el barco de Teseo” se refiere no a un barco en particular, sino a un *papel*, más como el papel de la nave insignia del almirante, que, en este caso, dicho papel puede ser desempeñado por diversos barcos en tiempos diferentes.

El tipo de ejemplo para el cual la teoría del mejor candidato no es en absoluto plausible es la identidad personal. En novelas y obras de ciencia ficción se discuten con frecuencia problemas relacionados con esto, como cuando la memoria y el carácter de una persona acaban en el cuerpo de otra, y se plantea la pregunta de si la persona con la memoria y el carácter de Smith (el continuador psicológico) o la persona con el cuerpo de Smith (el continuador corpóreo) es Smith. En este caso, la teoría del mejor candidato parece absurda. Desde fuera podría parecer que tenemos la opción de decidir si el cuerpo continuador era, desde luego, Smith quien había perdido sus recuerdos, etc., o bien decir que el continuador psicológico era Smith en un cuerpo diferente. Pero desde dentro —es decir, desde la perspectiva del Smith original—, difícilmente se tomaría como un asunto de decisión convencional. Si por ejemplo se le diera a este cuerpo (el cuerpo que tenía originalmente el Smith original) un trato desagradable después de tal cambio, ¿no se debe suponer que le lastimaría? Es difícil ver cómo la respuesta a esa pregunta puede depender de la decisión a la que puede llegar un observador bajo la teoría del mejor

candidato.

El convencionalismo acerca de la identidad es muy tentador, como vimos con el arbusto, y el regimiento, y el río, y el barco de Teseo discutido en la antigüedad. Pero confrontados con los dos barcos de Teseo de Hobbes, y, desde luego, con las amibas, vemos que el convencionalismo acerca de la identidad no es fácil en absoluto, y que estamos más comprometidos de lo que podemos suponer a dividir el mundo en determinadas líneas más que en otras. Cuando, más aún, se trata de mí en circunstancias similares, el convencionalismo parece perder totalmente su asidero. Aquí, el convencionalismo y la teoría del mejor candidato no nos hacen una invitación, sino que nos exigen que entendamos mejor nuestros propios pensamientos.

Este tipo de preguntas acerca de las personas, en un sentido metafísico, se refieren a qué o quién es una persona. Tales preguntas pueden estar relacionadas con la ética y la política en diversos aspectos. Pueden incidir, por ejemplo, sobre la ética y la política de la eutanasia. Pero hay otro tipo de pregunta ética o política que puede expresarse simplemente como “¿qué soy?” Este tipo de pregunta tiene que ver con la identidad de uno como una persona que pertenece a cierta familia, grupo o raza; son preguntas de identidad social. En estas conexiones, “identidad” tiene un sentido que, como en la historia de Nigel Dennis, se relaciona con un tipo o una cosa general. Una identidad gay o lesbica, una identidad indígena americana, o la de un lombardo en relación con la de un italiano, son todas cosas tipo, porque tal identidad es compartida. Desde luego, es particularmente importante el hecho de que sea compartida, y una insistencia sobre tal identidad es una insistencia en el modo en que es compartida. Una identidad, como ocurre en la novela de Nigel Dennis, también puede ser creada artificialmente. Por supuesto, no será construida por las intenciones de un club, sino por procesos sociales; y, nuevamente en común con el Club de Identidad de Dennis, alguna de las construcciones puede ser demostrablemente ficticia.

En este punto, es fácil decir que la identidad social es simplemente un estereotipo benigno autoaplicado, que es favorable, que brinda apoyo, y aplicado a uno mismo, más que uno que sea desfavorable y aplicado a nosotros por otros. Hay algo de verdad en esto, y nos recuerda cómo un estereotipo negativo puede ser convertido en una identidad positiva por la acción política. Pero esta simple explicación se basa en una diferencia fundamental. Un estereotipo asignado contra mí por otros incide en mí yo y penetra en él; es un obstáculo a mi libertad de vivir, de hecho o por convencimiento. Pero una identidad que yo tomo es una ayuda para vivir en los términos dictados por ella. La diferencia entre una identidad que es mía y que yo gustosamente reconozco como mía, y una identidad como la que alguien más simplemente supone que yo soy, constituye toda la diferencia del mundo.

Pero, ¿qué significa que un carácter general o papel o tipo constituya mi identidad? Aquí las relaciones entre el tipo y el individuo son cruciales. Es muy importante el hecho de que una identidad de esta clase no es mi identidad en el sentido particular. Si lo fuera, entonces, si diversas circunstancias me separaran de la vida y las lealtades que expresaban esa identidad, yo dejaría de existir. Más aún, si la forma de vida que

encarnaba esa identidad fuese destruida, las personas que la poseían cesarían de existir. Pero no es así, e insistir en esto no es simplemente un acto de pedantería o una afirmación de metafísica abstracta. El punto reside, más bien, en los pensamientos de las personas mismas que tienen tales identidades. Si esos desastres ocurren, las personas particulares todavía existirán, porque son ellas las que habrán sido dañadas o injustamente tratadas por este suceso. Si, por ejemplo, los indígenas americanos que residen en reservas son conscientes de la pérdida de identidad, son conscientes precisamente de *su propia* pérdida. Se dice frecuentemente que la destrucción de una cultura es una especie de genocidio; sin embargo, mientras que decirlo así tiene su razón de ser, es importante aclarar que saca de contexto la injusticia o el daño en el que se está incurriendo. La pérdida de una cultura puede ser considerada, desde un punto de vista conservador, como la pérdida de variedad o de una posibilidad humana; pero ése es un punto de vista externo, como lo es con la extinción de una especie.

Una parte esencial de la idea de identidad social es que un ser humano puede encontrar o perder identidad en los grupos sociales. Henri Tajfel, el fundador de la moderna teoría de la identidad social, definió la identidad social como “la conciencia de un individuo de que pertenece a ciertos grupos sociales unidos por valores o aspectos emocionales que todos comparten”.^[41] Esta explicación es suficiente para muchas de las preguntas que son investigadas en la psicología social bajo este título general, tales como las relaciones de la pertenencia a un grupo con la autoestima, las percepciones de prominencia, las relaciones entre grupos internos y grupos externos, etc. Sin embargo, es una definición demasiado amplia de lo que se podría llamar más estrictamente una *identidad*. Es así que alguien puede ser muy consciente de pertenecer a MENSA^[c] o a la *Royal Society*, y derivar autoestima a partir de ello, sin que ello constituya o contribuya poderosamente a su identidad. Desde luego, se puede pensar que una persona que encuentre su identidad fundamentalmente en ser parte de una de estas organizaciones se halla en un mal camino.

Una característica de la clasificación general o de tipo de lo que contribuye a crear la identidad de alguien es que es el pensamiento el que explica o provee la base a muchas de las actividades del individuo: emociones, reacciones o, en general, la vida. Es, desde el punto de vista de aquellos que la suscriben, una clasificación social profunda. Esto en sí mismo, por supuesto, no lo convierte en un concepto “clasificador”, es decir, un concepto fundamental para el conteo: el número de quebequenses presentes es el mismo que el de seres humanos presentes que son quebequenses. “Ser humano” es, a grandes rasgos, un término de la naturaleza, y “quebequense” un término cultural. Pero para quienes “quebequense” es un término poderoso de identidad social, es un término clasificador tan básico en la cultura como “ser humano” lo es en la naturaleza.

Sin embargo, es también típico de tales identidades que no son solamente análogas a las clasificaciones de la naturaleza, sino que están estrechamente relacionadas con la naturaleza. Por ejemplo, el haber *nacido* típicamente en una cierta relación con Quebec es lo que te hace ser un quebequense. Cuando personas de alguna minoría dejan de tratar de asimilarse, y optan por una identidad culturalmente distinta, buscan afirmar un

origen.

No siempre es tan simple como esto. Algunas personas pueden encontrar su identidad en una secta religiosa a la que voluntariamente se adhieren. Pero es típico en tales casos que tienen algún sentido de que esto no es simplemente optar por un grupo entre otros, sino que constituye encontrar algo que ya estaba allí; o regresar a casa —una especie de obediencia a la espléndida frase de Nietzsche de “conviértete en lo que eres”. En tal caso, aunque puedo sentir que he llegado ahí voluntariamente, aquello a lo que he llegado se halla fuera de mi voluntad: algo está dado, aunque debo escoger asumirlo. Esto es verdadero, por supuesto, también de aquellos otros casos, de identificación nacional o racial o tribal; se puede ejercitar la voluntad para llegar a coincidir con algo que ya soy de una manera que no se puede cambiar.

Esta es una razón, también, por la que hay una especial complejidad en relación con la identidad sexual. La adopción autoconsciente de una vida gay o heterosexual tiene su importancia, seguramente, porque no es meramente unirse a uno u otro club, sino que cuenta como el reconocimiento de algo. Al mismo tiempo, esa conciencia requiere también que el ser gay o heterosexual debiera ser no simplemente un asunto de determinismo genético o de desarrollo. Debe haber un espacio tanto para la naturaleza como para la voluntad.

Esto nos ayuda, quizá, para ver por qué la política de la identidad debiera ser tan esencial en nuestra vida ahora. Desde la Ilustración, una aspiración recurrente de la política distintivamente moderna ha sido por una vida que sea, desde luego, individual, particular, mía, dentro del alcance de mi voluntad, y al mismo tiempo exprese más que a mí, y le de forma a mi vida en términos que signifiquen algo porque se hallan más allá de la voluntad y que me están concretamente dados. Es la política, si te gusta, de la autorrealización. Ese término contiene en sí mismo obvias dificultades: es, incluso, gramaticalmente ambiguo entre actividad y pasividad, de una manera reveladora. Aquellas oscuridades son el producto no de una mera ineficiencia semántica, sino de tensiones políticas y personales no resueltas. Esta es una aplicación de una lección más general acerca de los problemas filosóficos de la identidad, que si encontramos sistemáticamente difícil saber qué decir, el problema se halla probablemente no en nuestras palabras, sino en nuestro mundo.

Segunda Parte

Ética

VI. La primacía de las disposiciones

I

Hay varias maneras de entender una búsqueda filosófica de lo que es primario o fundamental en la ética. Puede ser la búsqueda de la prioridad conceptual o de la dependencia, y un modo de entender tal prioridad sería en la forma de definiciones. Por un tiempo, Moore[42] pensó que lo *justo* había de ser definido como *productor del mayor bien*, y él entendió esto no como una definición concertada, sino como una explicación de lo que el término realmente significaba. Esto implicaba que una posición evidentemente contenciosa, como el utilitarismo generalizado o consecuencialismo, se encontraba no solamente en el lenguaje, sino también en su superficie. Como hipótesis lingüística, la sugerencia era muy poco satisfactoria y estaba expuesta al estilo de argumento de “pregunta abierta” que Moore mismo criticaba en otras definiciones sugeridas de *bueno*.

Hay un tipo diferente de pensamiento, que introduce un tipo diferente de prioridad, y es posible que esto haya sido lo que Moore con poca eficacia trató de expresar en términos de definición. Tales pensamientos tienen que ver con modos en que los diferentes tipos de valor, en particular el valor de diferentes tipos de cosa, pueden ser mejor entendidos, y pueden intentar establecer prioridades entre ellos. Pueden llamarse pensamientos “explicativos”. En el presente caso, de este pensamiento se desprendería que el valor de las acciones estaría derivado enteramente del valor de las cosas como están.

Ha habido una sugerencia lingüística precisamente en la dirección opuesta, de manera que *bueno* se debe definir en términos de *justo* o *debido*. Aquí, nuevamente, como hipótesis puramente lingüística, las propuestas no son satisfactorias en absoluto. Además es posible que algún otro pensamiento esté siendo expresado en esta forma y éste puede ser un pensamiento explicativo. Pero, ¿cuál sería el pensamiento explicativo en este caso? En el caso de la propuesta de Moore, sugerí que el pensamiento subyacente podría ser que las acciones derivaran enteramente su valor del valor de las cosas como están, pero no podemos simplemente tomar el camino opuesto con la presente propuesta. Nadie puede sugerir de una manera satisfactoria que las acciones en sí básicamente tengan valor, y que el valor de todo lo demás se derive del valor de las acciones. La acción tiene un fin, y las discusiones sobre la acción están generalmente dirigidas a asegurar algún bien ajeno a la acción misma. En discusiones sobre esto, frecuentemente (aunque no

siempre) hay quien quiere que su eventual acción acarree otra cosa, a la cual le adjudica valor. No puede ser verdad que el único valor básico sea el de las acciones. Si hay alguna concepción explicativa interesante que nos sugiera que *bueno* ha de ser definido en términos de *justo* o *debido*, no puede ser ésta.

Es difícil ver lo que esto podría ser. *Justo* y *debido* ciertamente expresan nociones que se aplican a acciones. Si el concepto dice, en cualquier forma, que estas nociones son básicas, mientras que nociones tales como *bueno* son derivadas de ellas, parece estar comprometido a decir que, en cualquier caso, todo *pensamiento* acerca del valor debe en algún sentido ser *pensamiento* acerca del valor de las acciones. ¿Cómo puede ser esto, si no es cierto —y obviamente no lo es— que todo valor se deriva del valor de las acciones?

Si *justo* o *debido* expresaran alguna noción moral específica tal como obligación, no parecería que puede haber una respuesta a esta pregunta. Ni siquiera todo pensamiento acerca del valor *moral* podría ser reducido a pensamiento acerca de la obligación. Esto está claro si uno considera que una situación moralmente mejor es aquella en la que las personas deben hacer lo que están obligadas a hacer. Las personas que creen en el valor moral básico de la obligación probablemente no se opondrán a este pensamiento, pero su contenido no puede ser interpretado enteramente en términos de obligación. En particular, aquellas personas no necesitan aceptar, y probablemente rechazarán, una aserción en el sentido de que cada uno de nosotros tiene la obligación de tratar de crear esa situación.

Sin embargo, *debido* no tiene por qué ser explicado en términos de obligación. Nos puede llevar a otra noción, más amplia: la noción de qué tiene uno más razón en hacer o promover. ¿Podemos explicar nuestros pensamientos acerca de lo que es valioso sólo en términos de lo que tenemos razón en promover? La respuesta parece ser una vez más “no”, puesto que nos encontramos con el mismo tipo de problema que enfrentamos antes, que frecuentemente pensamos que tenemos razón para promover alguna situación *porque* es valiosa.

Si hay una manera de evitar esta dificultad, tiene que encontrarse fuera de la teoría del valor. Tendremos que aceptar que uno puede tener una razón para promover algo sin pensar necesariamente que es valioso; así que uno puede tener razones para promover algo simplemente porque uno lo quiere, sin pensar que el obtenerlo sería valioso. Si queremos explicar la idea de lo valioso en términos de lo que uno tiene razón para promover, éste es el punto del cual tendremos que empezar. Podríamos decir, por ejemplo, que un aspecto es más valioso que otro precisamente en el caso de que más personas tuviesen razón para promoverlo. Otra sugerencia, no muy diferente en principio, podría ser que un aspecto valioso es aquel que un sujeto tiene razón para promover si tuviera el poder para hacerlo y que actuase solamente sobre la base de lo que la mayoría de las personas tienen las máximas razones para promover. Esto podría resultar en una versión utilitarista, reconocible como la teoría de Hare. Es una pregunta para investigar más adelante, qué variaciones sobre esta estructura podrían introducirse al interpretar más o menos generosamente la idea de lo que las personas *tienen razón* para

promover. Pero tiene que recordarse, desde luego, que si la estructura ha de servir al propósito que estamos discutiendo, la de explicar todo valor en términos de razones para la acción, las nociones de valor no deben darse por sentadas al establecer qué ha de contar como razones para la acción.

II

Hasta aquí me he movido muy fácilmente entre dos cuestiones diferentes: *dimensiones del valor*, tales como justicia, bondad, etcétera, y las *categorías* de cosas que están siendo evaluadas, tales como las acciones y las situaciones. Estos dos asuntos no están relacionados entre sí de una manera sencilla. Así que obviamente puede haber no solamente acciones que uno debería llevar a cabo, sino acciones buenas, y uno puede decir de su apariencia no sólo que son buenas, sino que deberían ocurrir. Así que uno no puede discutir prioridades entre las dimensiones de valor simplemente en los términos de las categorías de las cosas que están siendo valoradas. No quiero negar esto. Mi suposición hasta este punto ha sido tan sólo que aquellos que han querido hacer lo *debido*, o alguna noción similar, naturalmente comenzarán con las acciones, mientras que aquellos que ven lo *bueno* como noción básica preferirán, dada esta elección, ver como primordial la valoración de las circunstancias. Esto al parecer es bastante razonable. Una vez que se deja atrás la preocupación por las definiciones, de hecho, las preguntas más interesantes no tienen que ver tanto con las dimensiones del valor como con las categorías de elementos valorados, y me concentraré ahora en estas cuestiones.

Una estrategia reduccionista de categorías de elementos va a tratar de explicar el valor de todo en términos del valor de tan pocos tipos de elementos como sea posible — en el límite, en términos de un solo tipo—. Si es plausible cualquier visión monista, debe seguramente serlo en favor de la situación, porque sólo ésta cubre suficiente terreno. Sobre la visión monista, entonces, la cuestión del valor más fundamental —la llamaré la cuestión del valor— va a ser “¿Qué situación sería la mejor?”

El valor no es en sí mismo una cuestión práctica. *La cuestión práctica* (como la llamaré), lo básico para la gestión, debe ser “¿Qué he de hacer?” (No puede ser una pregunta impersonal como “¿Qué ha de hacerse?” La pregunta impersonal no será respondida como una pregunta práctica hasta que la pregunta en primera persona haya sido respondida por, o en nombre de, algún sujeto).[43] Así que debemos preguntar de qué manera están relacionadas la cuestión práctica y la cuestión del valor —la única pregunta de valor básica del monismo. Se han dado dos respuestas extremas a esto, así como algunas no tan extremas. Una respuesta extrema es que, en las ocasiones en que la acción ha de ser guiada por valores, la cuestión práctica debiera estar inmediatamente guiada por la cuestión del valor (esta es la respuesta utilitarista directa). La otra respuesta extrema es que los buenos resultados se promueven mejor si las personas, al considerar la cuestión práctica, raramente o nunca tratan de responder a la cuestión del valor.

La elección entre estas respuestas, o entre cualquiera de éstas y otras respuestas

menos extremas, no puede ser principalmente una pregunta filosófica. Concedidos los términos de la cuestión del valor, y dado algún entendimiento concreto de lo que lo hace valioso, debe ser en buena medida una cuestión empírica la de si los mejores resultados habrán de ocurrir si los sujetos se hallan generalmente dispuestos a plantear la cuestión del valor cuando deciden qué hacer. Este punto es reconocido por aquellos que argumentan en favor del utilitarismo indirecto al sostener que es empíricamente insatisfactorio que se obtengan mejores resultados si las personas tratan de responder a la cuestión práctica directamente en términos de la cuestión del valor.

Así que se requieren evaluaciones empíricas. Al hacerlas, debemos adoptar una visión externa de los sujetos, que vincule ciertos hechos acerca de ellos con los resultados que sus acciones probablemente han de producir. Necesitaremos contar con la información que probablemente adquieran y con su habilidad para valorar consecuencias con exactitud, pero también tendremos que conocer sus hábitos o disposiciones.

Esto no será simplemente un asunto de su inclinación o su falta de inclinación para responder la cuestión práctica en términos de la cuestión del valor. Tenemos que considerar las opciones que tienen para responder de dicho modo a la cuestión práctica. Si han de responder a la cuestión práctica en muchas ocasiones sin plantear la cuestión del valor, y (como espera el teórico) sus respuestas llevan a acciones que conduzcan a una mejor situación, entonces aquellos agentes deben tener algunas otras características en las que se puede confiar que expliquen cómo ocurre esto. Si pudiéramos confiar en una mano invisible o si, una vez más, estuviéramos preocupados simplemente por un sistema de recompensas y castigos externos, aquellas características podrían, en principio, consistir nada más en el interés propio elemental, algún patrón de esperanzas y temores. Pero no podemos creer en lo primero y nuestro interés no se halla en lo segundo. Estamos ocupándonos de un sistema ético o moral. Si hemos de considerar cualquier posibilidad que no sea la de los agentes respondiendo a la cuestión práctica directamente en términos de la cuestión del valor, tendremos que considerar otras y más determinantes características que pudieran poseer los agentes. Si los agentes han de generar situaciones deseables gracias al resultado de sus deliberaciones, pero las deliberaciones no son en sí mismas simplemente formuladas en aquellos términos, entonces los agentes necesitarán tener percepción de lo que es ético.

Hay una inestabilidad bien conocida en las estructuras de este tipo. Se requieren disposiciones éticas para generar situaciones deseables, pero no pueden permanecer simplemente como un mecanismo de caja negra para hacerlo. Necesariamente van más allá y proveen al agente un punto de vista a partir del cual puede valorarse el mundo. Desde ese punto de vista, no son simplemente las situaciones las que al parecer tienen valor. Esto se extiende al modo en que el agente entenderá, desde ese punto de vista, las disposiciones éticas mismas. Desde ese punto de vista, las disposiciones éticas parecerán tener un valor que va más allá de su capacidad de generar, mediante la acción, situaciones valiosas. Esta inestabilidad plantea una seria objeción a tales modelos. La objeción no necesariamente significa que los modelos sean inconsistentes. Éstos pueden ser, más bien, socialmente impracticables. Nuevamente, y algo muy interesante,

pueden no ser impracticables bajo todas las condiciones sociales posibles. Podría haber una sociedad, quizá de un carácter elitista, que realizase el modelo o una aproximación al mismo. La objeción será, más bien, que el modelo no puede realizarse en una sociedad que satisfaga ciertos requerimientos de transparencia y racionalidad autoconsciente: en particular, no podría realizarse en el tipo de sociedad que se supone satisfacer las expectativas bajo las cuales se proponen las teorías éticas de este tipo: publicación libre, investigación racional, crítica social abierta, etcétera.[44]

III

Hemos llegado a este punto considerando ideas de prioridad en la ética que primero fueron expresadas por medio de definiciones. Esa formulación fue insatisfactoria; se dio paso a la pretensión de encontrar un tipo de valor que pudiera ser básico y permitiera explicar otros en relación con él. Los “tipos” de valor incluidos aquí podrían entenderse o bien como dimensiones del valor (el bien comparado con la justicia, por ejemplo), o en términos de las categorías de elemento valorado. Las teorías reduccionistas se pueden aplicar a cualquiera de estos asuntos, o a ambos, y hay alguna correspondencia entre lo que se puede decir acerca de los dos, aunque dicha correspondencia no sea perfecta. Si uno considera teorías reduccionistas que abordan categorías de elementos valorados, la más plausible de ellas es la que hace de los resultados el elemento básico a ser valorado, y explica otros tipos de valoración en sus términos. Pero es una dificultad bien conocida que tan sólo ciertas hipótesis altamente probables hacen muy inestable el sistema que está basado en este concepto. Creo que aquellas hipótesis son verdaderas, que la teoría es inestable, y que esto es una objeción para aceptarla.

De hecho, hay varios tipos de cuestiones que podemos valorar en el pensamiento ético y hay un amplio rango —más amplio de lo que hasta el momento he sugerido— de conceptos en cuyos términos las podemos evaluar. He sugerido que una reducción a *situaciones* no es satisfactoria. Pero si ha de haber una reducción monista de los objetos de valoración ética, ésta es la mejor forma que podría asumir. Quedan las posibilidades de que podría haber una reducción de categorías menos detallada; o que las dimensiones de la evaluación podrían reducirse sin una reducción de sus objetos. Pero la conclusión más razonable parece ser que los intentos de reducir, en una u otra forma, están equivocados. Ya sea que uno considere la historia, la fenomenología, o las funciones sociales del pensamiento ético, no hay la presunción de que habrán de revelar prioridades generales de estos tipos, y ninguna razón para suponer que así será.

Hay algunas prioridades localizadas. Es verdad, hablando en general, que, en lo que se refiere a la justicia distributiva, la justicia de los resultados antecede a la justicia como una disposición de carácter: se tiene que explicar lo segundo en términos de lo primero. Con respecto a otras virtudes, hay otras cosas que decir: es necesario explicar la relación entre valentía como una disposición y la valentía de los actos valerosos, y la explicación no será la misma que en el caso de la justicia.[45] En tales casos, *hay* una presuposición

en favor de alguna relación de prioridad, en las tradicionales bases aristotélicas de que ciertamente no es un accidente, una cuestión de mera homonimia, que la misma palabra puede aplicarse a acciones y a personas, y debe haber algo revelador (aunque no necesariamente muy simple) que se pueda decir sobre las relaciones entre las diferentes aplicaciones de la palabra. Pero si planteamos alguna cuestión muy general de prioridades entre tipos de valor ético, no hay una razón particular para esperar una respuesta, y si no hay una razón particular, no hay razón. Ciertas acciones tienen valor porque conducen a situaciones valiosas: algunas porque expresan disposiciones valiosas; otras porque son las acciones que son, etcétera.

Esto no significa negar que puede haber explicaciones de nuestras concepciones éticas. Algunas de estas explicaciones son filosóficas. Así que se puede preguntar razonablemente por qué deberíamos tener un tipo especial de razón ética que llamamos una “obligación”, y la respuesta comenzará por considerar, de un modo filosófico y relativo a algunas suposiciones filosóficas acerca de la acción y la deliberación ética, qué es una obligación. Tendremos que considerar también qué es probable que resulte si tenemos la práctica de reconocer obligaciones como un tipo de consideración ética. En parte, aquellas respuestas serán empíricas, si bien no de una manera muy exigente, y pueden contribuir a una explicación de esta concepción ética. Se requiere, desde luego, una explicación si nuestra comprensión del pensamiento ético no resulta misteriosa (en el mejor sentido, naturalista).

Puesto que tales explicaciones incluyen una descripción de las obligaciones, hay otras maneras de entender las obligaciones con las cuales entrarán inevitablemente en conflicto. Puesto que buscan no hacer misteriosa nuestra conciencia ética, entrarán particularmente en conflicto con explicaciones de la obligación que la hagan esencialmente misteriosa. Puede ser que el entendimiento preteórico de la obligación la represente más bien misteriosa (aunque seguramente no tan misteriosa como aparece en algunas explicaciones teóricas de ella). Una explicación satisfactoria podría ayudarnos a entender la razón y valor de vivir una vida en la que las obligaciones cuenten como razones éticas pero, igualmente, las obligaciones podrían no resultar de la explicación con la resonancia que parecían poseer antes. Aunque es probable que, una vez que hayamos entendido y aceptado la explicación, podamos seguir viviendo una vida en la cual las obligaciones tengan algún papel; algunas personas podrían sentir que este papel no es suficiente y que la obligación ha muerto bajo la navaja explicativa. Hay de hecho algunos casos en los que el elemento que está siendo explicado no sobrevive porque no merece sobrevivir; la explicación enfoca la crítica sobre el elemento, más que servir para reivindicarlo o ayudarnos a aceptarlo. Alguien que aceptase la famosa declaración de Hume de “la virtud artificial de la modestia y la castidad”,^[46] por ejemplo, a duras penas supondría que dejó todo donde estaba (o al menos donde respetablemente se suponía que estaba).

Las ideas que uno puede tener de lo que es misterioso o no misterioso, por ejemplo, en relación con la obligación, serán influidas por supuesto por la imagen general que uno tenga de lo que son las convicciones éticas. Las explicaciones de nuestras concepciones éticas deben implicar explicaciones de los conceptos en los que aparecen, y que no pueden ser separadas de los asuntos epistemológicos, puesto que el conocimiento es en sí mismo una noción explicativa. En lo que concierne al pluralismo, sin embargo, la mayor parte de lo que se ha dicho hasta aquí debería ser compatible con una variedad de puntos de vista metaéticos acerca de la naturaleza y la epistemología de las convicciones éticas. El pluralismo que he recomendado no implica ninguna posición particular en ese nivel. Alguien con sólidos conceptos realistas podría, supongo, tener un motivo especial para plantear preguntas reduccionistas: contaría con una herramienta que le ayudase a encontrar qué tipos de hechos éticos hay. Pero esa motivación no va a alterar los fenómenos y la respuesta a su pregunta; si estoy en lo correcto, tendrá que ser que hay varios tipos de hecho ético.

Mientras que el pluralismo ético de este tipo no implica una epistemología más que otra, las implicaciones sí corren en la dirección opuesta. Nuestras conclusiones epistemológicas afectarán nuestro entendimiento reflexivo del pluralismo y del rango de elementos implicados en él. En particular, pueden dar lugar a la aceptación de cierto tipo de prioridad o basicidad, muy diferente de las que ya se han discutido.

Algunas posiciones metaéticas, incluyendo aquellas que son, para mí, las más plausibles, tienen la siguiente consecuencia: las características que las personas adquieren y ejercitan en la vida ética y que son distintivas de ella no se entienden de la mejor manera sobre el modelo de las capacidades cognitivas o de percepción, sino más bien sobre el modelo de las disposiciones de carácter. Llamemos a cualquier noción que tenga esta consecuencia una *noción de la disposición*. Es natural decir que las nociones de la disposición no son cognitivistas, pero esto es generalizar demasiado, pues en cualquier caso los no cognitivistas están comprometidos a sostener que no puede haber algo llamado conocimiento ético. Una noción de la disposición puede sostener que en cierta medida hay algún conocimiento ético, pero es conocimiento que se llega a adquirir y a ejercitar solamente al adquirir ciertas disposiciones de carácter.[47] Lo que se excluye por una noción de la disposición es el realismo acerca de los hechos o propiedades éticos; supongo que esto es excluido *ex hypothesi*, puesto que las únicas explicaciones del realismo ético que son inteligibles tratan de hacer que uno entienda las capacidades éticas básicas como análogos cercanos de la percepción sensorial o la intuición intelectual.[48]

Dada una noción de la disposición, se entiende que las disposiciones tienen cierto tipo de prioridad. Esto no es una prioridad *dentro* del pensamiento ético: no es más correcto de lo que fue antes decir que todo valor ético es el valor de las disposiciones. Lo que es, más bien, verdadero es que todo valor ético descansa *en* disposiciones. Las disposiciones son básicas porque la reproducción de la vida ética yace en la reproducción de las disposiciones. Ellas mismas se hallan entre los objetos de la evaluación ética, y son características en virtud de las cuales se piensa que las personas mismas son mejores o

peores; pero, de modo único entre aquellos objetos, hacen posible la evaluación de todos ellos. En un cierto sentido, dan el valor de aquellos otros objetos, aunque el valor de aquellos objetos no pueda ser reducido al de ellas.

He dicho que éstas son las disposiciones de los individuos, y es como las disposiciones de individuos que son replicadas y sustentan la vida ética. Es, por supuesto, también verdadero que una forma de vida ética es una cosa social y que incluye instituciones sociales, relaciones y papeles. Más aún, no hay razón para suponer que estos elementos sociales sean reducibles en ningún sentido a términos individuales. Nada en la concepción de disposición implica que debieran serlo. No es inconsecuente sostener al mismo tiempo que la vida ética es un elemento social con mucho contenido social irreducible, y que existe en la forma de disposiciones individuales. El segundo punto tiene que ver con la ubicación y realización de los pensamientos, deseos y actitudes relevantes, mientras que el primero concierne a su contenido y su explicación.[49]

Hay consecuencias prácticas importantes de la noción de la disposición, y de la prioridad que atribuye a las disposiciones. Si la vida ética ha de ser preservada, entonces estas disposiciones han de ser preservadas. Pero, de igual modo, si la vida ética que tenemos ha de ser efectivamente criticada y cambiada, entonces sólo puede hacerse de modos que puedan ser entendidos como una modificación apropiada de las disposiciones que tenemos. Desde luego, solamente una noción de la disposición, me parece, puede dar una explicación socialmente y psicológicamente realista de la crítica ética y sus efectos, una explicación que dé suficiente peso al hecho de que en realidad podemos explicar y entender la ocurrencia de actitudes éticas que encontramos prejuiciosas, limitadas, confusas, bárbaras, etc. Aquellas concepciones, por otro lado, que ven las características éticas básicas más como poderes puramente cognitivos, necesitan una teoría del error, y no la tienen. Sin una teoría convincente del error, están destinadas a encontrar que la crítica ética de gran escala es o bien imposible para ellas, o está condenada a ser puramente moralista.

VII. La estructura de la teoría de Hare

1

Espero que Dick Hare esté dispuesto a aceptar crítica con intención seria como una expresión de interés y respeto, pues es de hecho lo que siento al comentar extensamente su trabajo, quizá en exceso como posiblemente él, con toda razón, lo considerará. Tratando en un libro reciente (Williams, 1985)[*] de describir la filosofía moral y algunas importantes contribuciones modernas, encontré apropiado criticar sus concepciones en varios puntos del argumento. Aquí trataré de examinar la estructura de su teoría tal y como yo la entiendo, y en el transcurso de esto haré algunas de estas críticas nuevamente, pero de una forma distinta, respondiendo a lo que considero como la configuración total de su teoría. El diseño de este libro[**] le da a él la oportunidad de replicar y ésta es la razón por la que repito algunos de mis puntos de vista.

2

Considero que la estructura de la teoría de Hare[50] (permitiendo algunas formulaciones comprimidas) puede representarse del siguiente modo:

- a) el lenguaje moral tiene ciertas características necesarias.
- b) es necesariamente i) prescriptivo y ii) universal, y iii) estas características son suficientes para determinar la naturaleza del pensamiento moral a un nivel básico.

Debido a *bi* y *bii*,

- c) el pensamiento moral implica una identificación con las preferencias de todos,[51] y debido a *biii*, esto es todo lo que hay con respecto al pensamiento moral básico. Por lo tanto,
- d) el pensamiento moral básico es equivalente a una deliberación en cuasi-primer persona, gobernada por criterios de teoría de la decisión, sobre todo preferencias.

Es decir, el pensamiento moral básico es propiamente representado por lo que he llamado el *modelo del agente del mundo* (Williams, 1985:83ss). Pero

e) dentro del modelo del agente del mundo, pueden evaluarse elementos de diferentes niveles: acciones particulares, acciones tipo, políticas, principios, disposiciones, etcétera. En particular, sus criterios pueden aplicarse a la actividad de este mismo pensamiento moral básico.

Cuando son así aplicados, el resultado es que

f) el pensamiento moral cotidiano no debería tomar en su totalidad la forma de pensamiento moral básico (“crítico”); parte de él debería implicar una expresión “intuitiva” de disposiciones. El resultado es una teoría de dos niveles.

3

He presentado *f* como una consecuencia del resto de la teoría: inmediatamente, de *e* junto con algunos hechos supuestamente empíricos. Pienso que esto es correcto para la estructura lógica de la teoría, pero no revela adecuadamente la importancia de la parte desempeñada por la concepción de dos niveles al recomendar las primeras partes de la teoría. Obviamente *c* y *d* no son verdaderas. Más aún, incluso si alguna concepción tal como la de una *consideración imparcial* de todas las preferencias derivara del lenguaje moral o, independientemente de eso, se pensase que constituye el pensamiento moral básico, aún así podría pensarse que *d* y *c*, cuando se toman literalmente, implicaban una interpretación demasiado fuerte de esa concepción. Ante tales objeciones, la concepción de dos niveles ayuda a hacer aceptables *c* y *d*, salvando las apariencias que parecen hablar en contra de ellas. A varias características de la experiencia moral que sugieren una concepción diferente se les concede su lugar en el nivel “intuitivo”, mientras que *c* y *d* valen en el nivel básico o “crítico”.

Esta estrategia es similar a la que adoptó Sidgwick: tratar de mostrar que el utilitarismo, propiamente entendido, no tiene las consecuencias desagradables que parecía tener cuando se presentaba simplemente como una teoría de un solo nivel. Sin embargo, las dudas acerca de *c* y *d* tienen efectos más amplios. El hecho de que *c* y *d* no sean aceptados por muchos pensadores meticolosos hace no sólo oscuro sino inadmisibles que provengan de la naturaleza del lenguaje moral;[52] es decir, arroja dudas sobre el papel de *b*. Si la estrategia de dos niveles ayuda a defender *c* y *d*, ¿defiende con ello el papel de *b*? Aquí hay un contraste interesante entre la posición de Hare y la de Sidgwick. Sidgwick creía que *c* y *d*, o formulaciones similares, provenían de una intuición puramente racional para el efecto de que no puede haber una razón antecedente para preferir una parte de un bien asequible sobre cualquier otra parte, y él pensó que uno llegaba a ver esta verdad concentrando la propia mente sobre ella. Su teoría pretendía producir un equilibrio aceptable entre principios muy generales, por un lado, y reacciones morales espontáneas a situaciones familiares, por el otro. En particular, pretendía ayudar al teórico a justificar, y en ese sentido a entender, aquellas reacciones en términos de los principios generales;

iba también a ayudar a defender los principios en contra de la fuerza aparentemente contraria de las reacciones. Mientras que la epistemología de las intuiciones permanece irremediablemente oscura, Sidgwick pudo poner este proceso en alguna relación con la aprehensión que uno tiene de los principios básicos como él los concebía, diciendo que la teoría de los dos niveles, al dar lugar a experiencias que parecían hablar en contra de los principios básicos, ayudaba a concentrar la mente de uno en aquellos principios.

El uso, por parte de Hare, de la teoría de los dos niveles para hacer más aceptables *c* y *d* es muy parecido al de Sidgwick. Pero el uso que Hare hace de ella no tiene nada que ver con el modo en que deriva *c* y *d* en primer lugar. La teoría de los dos niveles no hace nada para hacer más plausible la aseveración de *que se deducen de la naturaleza del lenguaje moral* —es decir, para justificar el papel de *b*. Suponga que es verdad que si uno piensa mucho en *c* y *d*, con la ayuda de la teoría de los dos niveles, llegará a convencerse de ellas: eso no es suficiente. Es suficiente para Sidgwick, pero no para Hare, si es que ha de justificar el papel de *b*. Lo que debe ser verdad es que al pensar más intensamente *acerca del lenguaje moral*, uno se convencerá de ellas, y la teoría de los dos niveles no ofrece ninguna ayuda para eso.

Además de preguntar si las consideraciones acerca del lenguaje arrojarán este resultado más que cualquier otro, podemos preguntar por qué algo llamado “lenguaje moral” debería revelar algo, o algo confiable, acerca de estos asuntos después de todo. El lenguaje consiste en prácticas humanas; los seres humanos (como la teoría misma insiste) han padecido y padecen muchas ilusiones acerca de las relaciones de valor con el mundo, etc.; si el lenguaje puede plasmar o implicar cualesquiera proposiciones acerca de tales cosas, ¿por qué no debería plasmar ilusiones?

Sólo como una teoría moral, parecería que el sistema de Hare puede funcionar igualmente bien sin *a* y *b*. Podría proceder buscando lo que Rawls llama un “equilibrio reflexivo”, balanceando el material teórico general con las características de la experiencia moral, con la ayuda de la teoría de los dos niveles. De hecho, Hare ha resistido fuertemente este método, por razones a las que regresaré en 7.

4

La teoría de los dos niveles tiene sus propias dificultades. ¿En qué forma se suponen plasmar los dos pensamientos “crítico” e “intuitivo”? En la versión de utilitarismo indirecto de Sidgwick, los dos estilos de pensamiento corresponden a una distinción social entre dos clases de personas, la élite utilitarista y el resto. En esa versión —“utilitarismo de la Casa de Gobierno”, como pudiera llamarse— hay distinciones claras entre los pensamientos de las dos clases de personas. Por lo pronto, la distinción puramente lógica o semántica entre los pensamientos de primero y segundo orden se refleja en el hecho de que la élite piensa acerca de las prácticas de los otros, pero los otros no reflexionan sobre los pensamientos especiales de la élite —en particular (como lo deja en claro Sidgwick) porque no saben acerca de ellos. Más aún, los de la élite son

especiales no sólo por tener pensamientos de segundo orden; también piensan de un modo diferente acerca de asuntos prácticos particulares, puesto que la mayor parte de sus pensamientos acerca de estos asuntos expresan el utilitarismo directo, y pocos expresan principios de sentido común.

No es así como Hare representa la distinción entre los dos niveles de pensamiento. En otra parte (Williams, 1985:106-110) he dicho que considera a la distinción más bien como una entre dos periodos, el de la actividad práctica comparada con las horas propicias a la reflexión. Sin embargo, esto no es del todo correcto. Su distinción pretende, desde luego, ser psicológica más que social: recoge dos estilos de pensamiento, que bien pueden ser representados en la mente de una sola persona. Pero no recoge sólo dos momentos de pensamiento correspondientes a aquellos estilos, puesto que Hare supone que una persona puede pensar en ambos estilos al mismo tiempo. Más aún, él cree que hay en este aspecto varios tipos de personas, que se aproximan en diferentes grados a uno u otro de los dos arquetipos, reveladoramente llamados el “arcángel” y la “prole”. El primero de estos caracteres usa sólo pensamiento crítico —lo cual significa que sus poderes son tales que no necesita ningún otro. El segundo sólo usa pensamiento intuitivo, pues es incapaz de pensamiento crítico (*MT* capítulo 3, especialmente 3.1 a 3.3). Así que hay algunas personas —más como el arcángel, y más como la élite de Sidgwick— que pueden pensar en el nivel crítico la mayor parte del tiempo, y esto es visto como una capacidad valiosa, mientras que pensar en el nivel intuitivo es visto más bien como una necesidad y una imperfección. Cuánto tiempo es aconsejable pensar en una de estas formas, o en la otra, es una pregunta que Hare parece pensar depende de factores empíricos y carece de interés filosófico.

No creo que esta explicación sea coherente, o que los estilos de pensamiento que Hare busca describir puedan proveer lo que su teoría requiere. Es importante dejar en claro cuál es la objeción. No es una objeción (¿cómo podría serlo?) a todos los estilos de pensamiento reflexivo o de segundo orden, ni a la mera idea de que tales pensamientos pudieran estar entretajidos con otros en el transcurso de la práctica. Por supuesto que Hare está en lo correcto al decir que un general en acción puede brincar entre niveles táctico y estratégico de pensamiento (*MT* 52). Pero los pensamientos estratégico y táctico, a menos que el general se halle en un embrollo, no entran en conflicto, ni hay ningún conflicto entre las actividades del pensamiento en un estilo o el otro. Ni las reflexiones de la filosofía moral necesariamente tienen que entrar en conflicto con pensamientos prácticos de primer orden: aquellos del contractualismo, por ejemplo, no suelen hacerlo. La objeción es específicamente al tipo de teoría de Hare, la cual representa las respuestas intuitivas como profundamente arraigadas, rodeadas de fuertes emociones morales, lo suficientemente robustas para ver al agente a través de situaciones en que la reflexión sofisticada pudiera desviarlo, etc., y no obstante al mismo tiempo explica aquellas respuestas como un artificio para asegurar resultados utilitarios. La teoría ignora el hecho de que las respuestas no son sólo un mecanismo de caja negra para generar lo que es probablemente el mejor resultado bajo condiciones confusas. Más bien constituyen un modo de ver la situación; y tú no puedes combinar ver la situación de ese

modo, *desde* el punto de vista de aquellas disposiciones, y verla en el modo del arcángel, en el que todo lo que es importante es la máxima satisfacción de preferencias, y las disposiciones mismas son sólo un medio hacia ello.

Al decir que tú “no puedes combinar” estas dos cosas, no quiero decir que como un asunto de hecho psicológico sea imposiblemente difícil. Las personas, desde luego, tienen pensamientos que describen en estos términos —Hare mismo ha dicho que lo hace. El punto es que los pensamientos no son estables bajo la reflexión; en particular, tú no puedes pensar en estos términos si al mismo tiempo aplicas al proceso el tipo de reflexión exhaustiva que esta misma teoría propugna. Eso no es una mera aseveración psicológica. Es una aseveración filosófica, acerca de lo que está implicado en la reflexión efectiva y adecuada sobre estos estados particulares de la mente.

Podría decirse que la dificultad surge sólo porque la profundidad y fuerza de las disposiciones “intuitivas” han sido exageradas. Si las respuestas intuitivas son vistas sólo como presuntas reglas elementales, entonces la dificultad desaparece, o al menos no es tan severa.[53] Eso es correcto, pero la posición que provee no es la de Hare. Más aún, no le permitiría a Hare hacer todo el uso que hace de la teoría de los dos niveles. Al igual que Sidgwick, como ya dije antes, usa la teoría para salvar las apariencias, explicando las reacciones ordinarias de las personas en relación con los principios utilitarios básicos. Pero cuanto más representa la teoría las reacciones intuitivas como meramente superficiales, provisionales e instrumentales, menos apariencias salva: no explica lo que las personas sienten y piensan, sino que sugiere otra cosa en la misma área de lo que ellos podrían útilmente sentir y pensar.

5

Es importante recalcar que la teoría de los dos niveles en su forma presente no es la única solución, dado el resto del material, al problema de lo que debiera ser el pensamiento moral cotidiano. Su postura lógica, lo sugerí antes, es que está apoyada por la posición básica junto con algo de material empírico y, si hay dificultades con ella, entonces puede resultar que alguna otra solución esté mejor apoyada. Si el pensamiento moral básico es definido en términos arcangelicales, entonces se necesita *algún* suplente en la práctica cotidiana, puesto que el pensamiento puramente arcangelical requiere una cantidad indefinida de conocimiento y poder deliberativo (es de lo más estricto porque Hare interpreta de una manera muy ambiciosa la idea de estar imparcialmente ocupado con las preferencias de todos). No obstante, Hare mismo no puede ver el pensamiento crítico, no intuitivo, exclusivamente en términos arcangelicales, puesto que piensa que es algo que la mayoría de nosotros puede conducir a veces y, cuando es visto en esa luz, desde luego se convierte en una cuestión empírica hasta qué punto se requiere un sustituto, como él lo sostiene. Pero entonces el sustituto debe también ser la cuestión empírica.

Es obvio (como desde luego lo es desde la historia del utilitarismo) que hay muchos

candidatos para ser sustituto, muchos tipos de pensamiento moral cotidiano que podrían en la práctica producir los mejores resultados. Disposiciones profundas, que Hare favorece y que producen sus propios problemas especiales, son solamente una. Hay reglas aceptadas, utilitarismo directo aplicado a una membresía limitada de beneficiarios, y otras. Procediendo hacia abajo desde la explicación del pensamiento moral básico, y tomando en consideración los hechos empíricos apropiados acerca de los efectos en la práctica de diferentes tipos de práctica moral, uno podría llegar a cualquiera de estas soluciones. El hecho de que Hare arribe a este particular sustituto, en vez de a otro, puede no ser simplemente el producto de la creencia empírica. Puede ser que sea atraído a esta solución precisamente porque es, al menos, *prima facie* mejor que las otras para salvar las apariencias de la experiencia moral. Pero se necesita más argumentación para demostrar que la posición a la que uno llega procediendo hacia abajo desde la cima del sistema, y la posición que mejor salva las apariencias de la experiencia moral, es probablemente una y la misma. Si se adopta el método del equilibrio reflexivo entonces, desde luego, hay una razón para salvar las apariencias —esa es parte de la idea. Pero si el método es el de descender del lenguaje moral a los principios básicos, y de ellos, a la luz de los hechos empíricos, a las formas más deseables de pensamiento moral cotidiano, simplemente no hay garantía de que las apariencias merezcan ser salvadas.

6

Planteé (en 4) una dificultad con la teoría de los dos niveles en su forma presente, que favorece las disposiciones profundas. La dificultad es, en resumen, que uno podría no pensar en el nivel cotidiano “intuitivo” del modo que la teoría requiere mientras uno está plenamente consciente de lo que está haciendo: en particular, mientras uno entienda completamente en términos de la teoría misma lo que uno está haciendo. Esto, si es real, es una dificultad; pero ¿por qué exactamente esto constituye una objeción a la teoría de Hare? Podría ser vista como algo que no plantea una objeción, sino que muestra sólo algo acerca de la teoría, que no puede funcionar si las personas, conscientes en todo momento, la creen. Tales propiedades de las teorías éticas, en particular de las teorías utilitaristas, han sido estudiadas por Parfit[54] (Parfit, 1984, parte I), quien argumenta que no necesariamente significa que tales teorías debieran ser rechazadas. Pero esto deja problemas de quién ha de aceptar tales teorías, y en qué espíritu; y si no es posible que cualesquiera, o muchas personas debieran aceptarlas, cuál es entonces el estatus de la teoría, y el propósito del teórico al anunciarla.

En el caso del utilitarismo de la Casa de Gobierno, la respuesta a estas preguntas es directa. La teoría está dirigida a la élite, y no hay problema acerca de que la élite la crea plenamente, todo el tiempo. El problema difícil es político, de cómo gobernar una sociedad moderna que está controlada por una élite tal —en particular, gobernarla decentemente, de acuerdo con valores que podría esperarse que la élite pudiera (de otra manera) suscribir—. La teoría de Hare, sin embargo, está dirigida a personas de las que

se espera tengan disposiciones que no pueden ejercitar en el modo que requiere la teoría y al mismo tiempo entenderla en los términos que la teoría provee. Si se les pide que atiendan esta teoría se les está pidiendo también, en un grado significativo, que la olviden. Podría haber una teoría que pidiese ser tratada de ese modo, pero es obvio que no es el caso con la teoría de Hare. Por el contrario, esto pretende apelar a y sustentar un espíritu de reflexión crítica y de autoexamen, y Hare deja en claro que él pretende que su filosofía contribuya a la mejoría moral a través de la autocrítica. Debido a que la teoría tiene estas aspiraciones, la dificultad también es una objeción.

7

En su primera teoría, Hare separó su descripción del lenguaje moral de cualquier determinación del contenido de los principios morales. Él también estaba dispuesto a ver esa separación como una aplicación de la distinción hecho-valor —incluso, más que eso, como una expresión primordial de ella—. La filosofía misma tenía que estar en el lado “realista” de la distinción: tanto porque eso estaba de acuerdo con una imagen general de la filosofía como análisis lingüístico, como, en particular, porque el anuncio por parte de la filosofía de la distinción hecho-valor carecería de la fuerza crítica que se supone tener si fuera en sí misma una expresión de valor.

En la segunda teoría que estoy considerando, se piensa que la descripción del lenguaje moral, junto con información empírica, determina el contenido de los principios morales; y el grado en el que la descripción del lenguaje moral determina la naturaleza del pensamiento moral, incluso sin información empírica, ciertamente sirve ya para eliminar alternativas sustanciales. (Hare dice que lo que es determinado en este nivel es la *forma* del pensamiento moral (básico), pero entonces muchas preguntas sustantivas acerca de él deberían ser acerca de su forma). Debido a que Hare asoció sus primeras formulaciones tan estrechamente con la expresión de la distinción hecho-valor, estos desarrollos han llevado a algunos críticos a decir que Hare ha cambiado de parecer acerca de esa distinción. Sea como sea —pienso, de hecho, que Hare puede responder a esta afirmación—, lo relevante aquí es que no ha cambiado de opinión sobre lo que es la filosofía, y que tiene una concepción más fuerte ahora que antes acerca de sus poderes.

Hare siempre ha querido, y de una manera honesta, que la filosofía moral tenga un efecto práctico, que haga una diferencia. Él no considera un obstáculo que su concepto de la filosofía moral sea, a grandes rasgos, una rama de la lógica filosófica. Por el contrario, la filosofía moral puede hacer una diferencia solamente porque tiene autoridad, y puede tener autoridad solamente debido a su situación neutral como un tema de lógica o de lingüística. Siempre ha sostenido esto de alguna forma, pero su presente teoría da una explicación muy especial de qué es esta autoridad. Lo crucial ahora es la creencia de que esta materia neutral pueda aportar fundamentos.

Es esta creencia la que explica su burla de aquellos que usan el método de buscar un equilibrio reflexivo e inician su filosofía moral con las “intuiciones morales” de la gente.

Hare no objeta en principio esto, como lo hacen algunos utilitaristas, porque estas intuiciones sean demasiado conservadoras; los principios “intuitivos” que él pone al final de su investigación tienden, en algunas áreas, a ser más conservadores que aquellos que hubieran sido extraídos en primer lugar. De modo similar, la objeción no puede ser simplemente que su método no sea lo suficientemente crítico o radical, si esto se toma en algún sentido general: en un sentido general algunos de estos filósofos son notablemente críticos o radicales. Su objeción es que, ya sean conservadores o radicales en el resultado, las intuiciones y las teorías extraídas de ellas por esos filósofos son meros prejuicios: lo que significa que carecen de fundamento. Esto provee un sentido en el que piensa que estas investigaciones no son lo suficientemente críticas o radicales —el sentido cartesiano de que no han llegado a los fundamentos.

La concepción presente de la filosofía moral de Hare y su relación con la práctica tienen, entonces, varias hebras, con diferentes historias. Él cree ahora:

- i*) La filosofía moral es (a grandes rasgos) una materia lógica o lingüística;
- ii*) provee fundamentos;
- iii*) nos ayuda a reflexionar claramente sobre nuestros pensamientos morales, y en particular —debido a *i*— a pensar acerca de lo que queremos decir;
- iv*) cuando lo hacemos, descubrimos *ii*.

Hare siempre ha creído *i*. No siempre ha creído, como ahora lo cree, *ii* y *iv*. Sin embargo, él cree ahora que en alguna parte la autoridad de la filosofía moral se deriva de *ii*; como hemos visto, aquí descansa su rechazo de otros métodos.

Discutí antes la dificultad acerca de la teoría de los dos niveles y sugerí, en 6, que la razón por la que esa dificultad era una objeción a Hare se hallaba en la concepción que adoptó de la filosofía. En términos de este esquema, es una objeción referente a *iii*. Hare siempre ha aceptado *iii*; desde luego, antes de que llegara a creer *ii*, tuvo que hacer todo el trabajo de explicar por qué la filosofía, como un estudio neutral del lenguaje, tenía autoridad y relevancia para la vida práctica. En sí mismo, *iii* no presenta ninguna dificultad, por lo menos de este tipo. Otros pueden estar menos inclinados de lo que Hare lo está a enfatizar su aspecto puramente lingüístico, pero todos pueden reconocerlo como una declaración digna del impulso socrático, una expresión de los valores de la transparencia social y personal. Sin embargo, cuando Hare pasó a creer *ii* y *iv*, surgió la posibilidad de que *iii* entrase en conflicto con las conclusiones de *i*. Iba a ser una pregunta abierta la de si los fundamentos que *i* ahora proveía eran tales como que la práctica recomendada por *iii* pudiera ser completamente llevada a cabo. Hare siempre asumió que *iii* era una proposición obvia de las metas de la filosofía, y también que los resultados de la filosofía, como el tema objetivo descrito en *i*, pudiera ser consistentemente, y sin lugar a dudas de utilidad, conocido por cualquiera. Él ahora no tiene manera de asegurarse de que estas dos cosas —la concepción externa de lo que es la moralidad y la representación interna de ella en la práctica moral— necesariamente habrán de acoplarse.

Otras objeciones que he mencionado dan por resultado un rechazo a la idea de que una investigación lingüística habrá de producir fundamentos. No discutiré aquí la pregunta de si debería esperarse que la moralidad deba tener fundamentos, ni la paradoja de que Hare (y él no está solo en esto) debería cambiar a una concepción fundacionista de la moralidad mientras que la filosofía ha estado adoptando una concepción menos fundacionista de todo lo demás. El punto en el que estamos es solamente acerca del método. Ya he mencionado la razón básica de Hare para pensar que la filosofía no debería proceder por reflexión (en primer lugar, al menos) sobre nuestras “intuiciones”, opiniones o experiencias morales. Su razón es que ellas son sólo nuestras: como una objeción, descansa en su creencia en los fundamentos. Su particular insistencia en una investigación lingüística, como contrapartida a tales reflexiones, descansa en alguna parte del punto general *i*; pero la forma especial adoptada por su investigación lingüística revela que en sí misma está condicionada por la búsqueda de fundamentos, y por el deseo de escapar de lo que es sólo “nuestro”. [55]

Para aquellos que no están de acuerdo con Hare en que un método lingüístico puede aportar fundamentos, él no da ninguna razón puramente metodológica de por qué no se debería empezar con “intuiciones” —es decir, una razón independiente del punto sustancial sobre el cual él y ellos están en desacuerdo—. Ellos mismos pueden creer en fundamentos, pero basados en que el modo de encontrarlos es empezando con nuestras opiniones morales. O pueden no creer en fundamentos en absoluto, y estar sólo ocupados con las implicaciones, presuposiciones e incoherencias de aquellas opiniones. En relación con cualesquiera de estos grupos, Hare se halla en una situación familiar en la filosofía (era la misma situación de Descartes) de que su objeción a su método descansa solamente en lo que él cree que son los resultados actuales del suyo propio.

Sólo porque las conclusiones de Hare gobiernan su método, y sus conclusiones son tan fundacionistas, él responde claramente en sus propios términos una pregunta que otros filósofos con frecuencia simplemente ignoran. ¿Por qué debiera esperarse que la reflexión crítica con la que la filosofía moral está comprometida resultara en una teoría ética? Hare tiene una concepción de la filosofía moral y sus metas que surgen naturalmente en tal teoría. Aquellos que rechazan esa concepción, pero aún así buscan una teoría para sistematizar nuestras opiniones morales, le deben tanto a él como al resto de nosotros una explicación de por qué esperan que nuestro mejor entendimiento de nuestra vida ética asuma tal forma.

VIII. Subjetivismo y tolerancia

Bertrand Russell dijo más de una vez que le incomodaba un conflicto, como él lo veía, entre dos cosas: la fortaleza de la convicción con la que había sostenido sus creencias éticas, y las opiniones filosóficas que tenía acerca de la naturaleza de aquellas creencias éticas; opiniones que eran no cognitivistas, y en algún sentido subjetivistas. Russell sentía que, de alguna manera, si él no pensaba que sus creencias éticas eran objetivas, no tenía derecho a sostenerlas de una manera tan apasionada. Esta incomodidad no era algo que Ayer comentara o discutiera en su explicación de la filosofía moral y opiniones éticas de Russell, al menos en el libro que escribió para la serie *Modern Masters (RS)*.^[d] Quizá esto se debió a que no era una incomodidad que Ayer mismo sintiera. Sus propias concepciones filosóficas acerca de los fundamentos de la ética fueron en todos los periodos, en cualquier caso, no cognitivistas, y pienso que a él no le importaba que se les llamara “subjetivistas”. Él, desde luego, argumentó que la supuesta diferencia entre el objetivismo y el subjetivismo en la ética no tenía validez, y que los filósofos que se consideraban objetivistas no podían lograr más que aquellos que admitían que eran subjetivistas. Ayer basó esto principalmente en la idea de que las afirmaciones hechas por los objetivistas en favor de la objetividad, la verdad objetiva, y demás juicios morales, no agregaban nada a estos juicios; en lo que concernía a las conclusiones morales, el objetivista estaba diciendo lo mismo que el subjetivista pero con voz más estridente.

De esta manera, mientras que él pensó que las pretensiones adicionales del objetivismo no hacían ningún trabajo real, Ayer no concluyó que la distinción entre el subjetivismo y el objetivismo careciese en absoluto de contenido. No rechazó la distinción completamente, porque pensó que había una concepción metaética que él sostuvo y que cualquier objetivista rechazaría, a saber, que las expresiones morales no formulaban hechos, ni eran (realmente) verdaderas o falsas. Desde luego, al decir que ningún trabajo real se lograba por afirmaciones de que las expresiones morales pudiesen establecer hechos, Ayer quedó en desacuerdo con un objetivista, así que tenía que haber alguna diferencia entre su propia posición y la del objetivista. La afirmación de vacuidad no era en sí misma vacía.

Ayer admitía que muchas expresiones morales —la afirmación de que alguien es un cobarde, por ejemplo— tenían lo que en la concepción de cualquiera sería un componente fáctico: éste, por un bien conocido estilo de análisis, se suponía que era separado de la dimensión distintivamente moral de la expresión. Él admitió, más aún, que en el lenguaje ordinario las expresiones morales eran frecuentemente llamadas

“verdaderas” o “falsas”: y en esto él vio un prejuicio comprensible. Sin embargo, hay un problema con los demás elementos que él suponía necesarios para que éstas fueran *realmente* verdaderas o falsas; es decir, con lo que, de acuerdo con él, el objetivista creía y él no. Naturalmente aquí se invocaba una versión diluida del principio de verificación: si las afirmaciones morales fueran realmente verdaderas o falsas, tendría que haber algunos procedimientos establecidos, inherentes al significado de aquellas afirmaciones, para comprobarlas. Esto es lo que él tenía en mente cuando atacó el objetivismo en la forma del intuicionismo que estaba relacionado con el Oxford de sus primeros días, en particular con Prichard y Ross. Debido a que la teoría del intuicionista no aportaba alguna forma de verificar las supuestas verdades de la moralidad, Ayer sostuvo que una afirmación de la que pudiésemos tener conocimiento de dichas verdades o creencia razonada en ellas era falsa, y de esto él quizá concluyó que era igualmente falso sostener que fuesen en absoluto verdades. Puesto que Ayer estuvo de acuerdo en que uno podía, desde luego, sostener concepciones morales, se hallaba implícito en su posición, como en cualquier otra que niegue valor de verdad a las afirmaciones morales pero no las rechaza del todo, que uno puede tener una posición moral (asentir a ella, etcétera) sin sostener que es verdadera. Si la actitud hacia las proposiciones morales que uno sostiene puede ser llamada “creencia”, entonces hay proposiciones en las que uno puede creer sin creer que sean verdaderas.

Ayer rechazó esta conclusión al negar que una actitud favorable para las proposiciones morales fuese, propiamente hablando, una creencia. Él prefería expresar su posición sobre esto en términos del acto de hablar; la afirmación era que el que hablaba de moral no estaba formulando hechos, sino otra cosa. En *Language, Truth and Logic* el candidato para el otro acto de hablar era “expresar los propios sentimientos”; hacia 1949 eran cosas tales como prescribir, dar permiso, mostrarse uno como favorablemente dispuesto, expresar una resolución, etc. Desde luego, desde al menos 1949 en adelante Ayer parece haber adoptado una versión simple del prescriptivismo, el tipo de teoría que R. M. Hare iba a expresar en una forma más compleja en *The Language of Morals*.

Esta atracción a un acto de hablar alternativo puede proveer una fórmula para evitar las dificultades en torno a la creencia y la verdad, pero no provee (como Ayer y otros quizá creyeron) una forma independiente de entender la aseveración de que las expresiones morales no formulan hechos o de que tengan un valor de verdad. Las personas, cuando hacen observaciones de índole moral, por lo regular no se ven sólo como algo que simplemente prescribe, expresa una resolución, etc.; esto puede ilustrarse mediante su uso ordinario de “verdadero” y “falso”, que se mencionó antes, y suelen pensar en otros solamente cuando consideran que las afirmaciones de éstos carecen de base. La teoría filosófica de que lo que la gente está (realmente) haciendo es (solamente) prescribir, expresar resoluciones, etc., obtiene la fuerza que pueda tener del concepto, independientemente apoyado, de que no pueden estar expresando hechos. Desde luego, la fenomenología del pensamiento moral es en algunos aspectos especialmente reacia a una interpretación prescriptivista, un punto al cual regresaré.

Ayer introdujo en 1949 la idea de que tener razones para una prescripción moral

supuestamente iba a hacer posible abarcar un rango de comentarios morales más amplio que el que lograba el emotivismo. Como lo dijo él:[56]

Al decir que Bruto o Raskolnikov actuaron correctamente, me estoy dando a mí mismo y a otros permiso para imitarlos si surgiesen circunstancias similares... De un modo similar, al decir que actuaron incorrectamente, expreso una resolución de no imitarlos, y me propongo también disuadir a otros. Puede pensarse que el solo uso de la palabra despectiva “erróneamente” no es tan disuasiva, aunque tiene alguna fuerza emotiva. Pero es aquí donde entran las razones. Disuado a otros, o en cualquier caso esa es mi intención, al decirles por qué pienso que dicha acción es errónea.

Hay varias inconsistencias en este pasaje; por ejemplo, la idea de que de mí depende dar a alguien permiso para imitar a estos personajes. Más generalmente, hay algo que causa extrañeza en la idea de que los razonamientos que ofrece podrían ser suficientes para explicar “por qué pienso que dicha acción es errónea”. Lo que Ayer dice no es eso; simplemente explica cuándo puedo modificar la actitud de alguien más. Las conexiones entre tener una razón, dar una razón y tratar de persuadir a alguien, son mucho más complejas que lo que esto implica. En parte, este es un problema que quedó del emotivismo y, en parte, un problema compartido por el prescriptivismo.

Fue característico de la perspectiva de Ayer que el cambio al prescriptivismo no hiciera mucho más que sugerir nuevos modos de persuasión. Hubiera hecho más alertando a Ayer hacia una idea diferente de objetividad, asociada con la posibilidad de que las proposiciones morales pudiesen expresar objetivamente prescripciones universales (como en el trabajo posterior de Hare, y de una forma diferente en Kant). Tal concepción, por supuesto, separa la cuestión de la objetividad de la cuestión de si el discurso moral es como tal formulador de hechos: en la expresión kantiana, los principios morales correctos y su fundamento, el imperativo categórico, son *objetivos*, porque supuestamente están fundamentados en los requerimientos de la razón práctica, pero no formulan hechos, porque son imperativos. La objetividad es la de una construcción, no la de un descubrimiento. Ayer nunca consideró tal posibilidad. Para él, en línea con las preocupaciones positivistas, los asuntos de objetividad y de (a grandes rasgos) cognitivismo eran siempre los mismos.

Ayer insistió en que sus análisis metaéticos no tenían implicaciones para el pensamiento moral de primer orden. Su argumento para esto fue una versión prescriptivista aceptada de la diferenciación hecho-valor, junto con la afirmación de que el análisis metaético, por ser análisis, es descriptivo. A Ayer le preocupaba poner énfasis en el punto de que el análisis subjetivista no lleva a la conclusión nihilista de que las consideraciones morales son triviales o carentes de importancia. Esto fue lo más que se acercó a identificarse con, si bien no a compartir, la preocupación de Russell, y quizá aclaró su punto de vista debido a críticas provenientes de los cristianos Cyril Joad y otros, que tendían a acusar al positivismo de frivolidad moral o de cosas aun peores. Igualmente, Ayer no supuso que el análisis metaético llegase a tener implicaciones en la dirección opuesta; no pensaba que las consideraciones subjetivistas pudieran ser usadas para apoyar otras conclusiones más positivas, como la tolerancia. Esta es otra diferencia respecto de Russell, quien, en un conflicto enteramente comprensible con su

preocupación acerca de subjetivismo y la fortaleza de sus convicciones, estaba dispuesto a coincidir con J. S. Mill en que los fanáticos eran fortalecidos al menos por algunas variantes del cognitivismo.

Por supuesto, aun concediendo las suposiciones de Ayer, no se nos dice qué es lo que la filosofía debería hacer: las mismas suposiciones nos muestran por qué no se nos puede dar una respuesta. Si hemos de llegar a la conclusión, como Ayer y muchos otros partidarios de la distinción hecho-valor lo hicieron, de que la filosofía no puede ocuparse de la moralidad, necesitamos una premisa adicional, que Ayer aceptó, de que la filosofía está confinada al análisis. Más aún, la idea de que la filosofía no podía apoyar conclusiones distintivamente morales se fortaleció con la creencia de que “apoya” tenía que significar una de dos cosas: o bien “implica” o, bien, alternativamente, “por una ley estadística anima a las personas a pensar...” Ayer aceptó también esto, otro legado positivista. Eso dejó sin responder si profesar ciertas concepciones filosóficas pudiera motivar, como hecho estadístico, ciertas actitudes morales o políticas, y en este sentido empírico “apoyarlas”. Pero Ayer también al parecer pensó que, si esto resultaba ser así, no significaba que el filósofo que profesaba las concepciones tenía responsabilidad en esa dirección. El malentendido era, por así decirlo, culpa de quien escuchaba. Hay algo inocente, o (en una concepción más severa) convencionalmente académico, en esta perspectiva. El contenido de las expresiones morales es macadamamente reducido a su fuerza, y su fuerza virtualmente reducida al efecto que produce. Pero las observaciones de la filosofía, los resultados del análisis, permanecen seguros en su contenido conceptual, y significan ni más ni menos que lo que dicen. El efecto que puedan tener es considerado como algo totalmente independiente de ellos.

Dados estos conceptos, fue una especie de aberración cuando, en la conferencia Eleanor Rathbone que dio en 1965,^[57] Ayer explicó en términos históricos y políticos el hecho de que los filósofos franceses habían ofrecido más conceptos (morales) acerca de la política que los filósofos británicos. No dijo simplemente, como él debió haber pensado que tenía el derecho a decir, que los filósofos británicos se apegaban a la filosofía. No obstante, permaneció fiel a la separación que había aceptado entre filosofía y concepciones políticas de primer orden, y claramente la estableció con exageración surrealista en esta conferencia. En la primera parte de la misma, recorrió a través de un ensayo tutelar trece razones que han sido dadas explicando por qué debiera uno obedecer la ley. Entonces, procedió, poco filosóficamente, a describir su postura personal, veladamente liberal:

En este asunto soy como los demás; no tengo nada nuevo que ofrecer. Sólo los viejos y conocidos principios liberales; viejos, pero no tan firmemente establecidos como para que podamos considerarlos ya nuestros. El gobierno representativo, el sufragio universal, la libertad de palabra, la libertad de prensa, el derecho al contrato colectivo, la igualdad ante la ley, y todo lo que va con el así llamado estado del bienestar. No es una bebida embriagadora. En los tiempos que corren, estos principios sólo crean alboroto y entusiasmo cuando son amenazados. Para la mayoría de nosotros la participación en política asume la forma de protesta; protesta contra la guerra, contra las acciones agresivas de las potencias principales, contra el maltrato de los prisioneros políticos, contra la censura, contra el castigo capital o corporal, contra la persecución de los homosexuales, contra la discriminación racial; todavía hay mucho contra lo cual hay que estar. Sería más

romántico estar marchando adelante hombro con hombro bajo una brillante nueva bandera hacia un magnífico nuevo mundo. Pero no lo sé: quizá es el efecto de la edad. Realmente no siento la necesidad de nada para reemplazar este radicalismo principalmente utilitarista, principalmente tolerante y no dramático.

En este pasaje es injusto consigo mismo en más de una manera. Es injusto con el vigor con el cual él, desde luego, hizo campaña en favor de estas causas; es injusto al implicar que a los 55, la edad que tenía entonces, él se hallaba algo cerca del fin de sus energías o su compromiso con aquellas campañas. Pero quizá la implicación más emotiva de la conferencia en su conjunto es la sugerencia, contraria al tenor de lo que creía en filosofía, de que si la política fuese más ideológica, la filosofía política sería más vívida de lo que era; más vívida, desde luego, que la demostración que él mismo había dado de ella.

Otra consecuencia, insistía Ayer, que no se derivaba de su concepción metaética, era el relativismo. En particular, la concepción metaética no podía implicar un tipo específico de actitud relativista, que estuviera caracterizada por la tolerancia de prácticas morales divergentes. Si al hacer declaraciones morales prescribo, entonces, desde luego, prescribo para todos. Debe haber algo de verdad en esto; la tolerancia es una actitud sustancial y ciertamente no va a surgir de un análisis del lenguaje moral. Sin embargo, al mismo tiempo, muchas personas han pensado, como (lo mencioné antes) Russell pensó, que el subjetivismo podría tener algo que ver con la tolerancia: que había algo más general para ser aprendido del hecho de que algunos fanáticos son fuertemente objetivistas. Quizá si relajamos algo la noción de “consecuencia”, y permitimos un tipo de reflexión algo más amplio que el que permitía Ayer, puede ser una pregunta real la de si la conciencia del subjetivismo no debiera tener algunas consecuencias para la tolerancia y, en esos términos, para el liberalismo. Esta es la pregunta de la que me ocuparé en el resto de este artículo.

Algunas concepciones cognitivistas de hecho proveen motivos adicionales al fanatismo; por ejemplo, al mantener esperanzas en una recompensa divina, dando un aliento divino al fervor, ofreciendo certezas de que los ignorantes están siendo ayudados, etc. Una filosofía más escéptica podría eliminar algunas de estas incitaciones, un punto de vista muy común por lo menos desde Montaigne, y que es característico de la Ilustración. Ayer no le dio importancia a este punto de vista, quizá porque pensó que si el contenido cognitivista era identificablemente religioso, entonces las concepciones serían “fácticas” y no morales. El único elemento moral, en su análisis, será la prescripción de obedecer a Dios o a cualquier cosa que éste pueda ser. Así que podría decirse, si bien de manera estrecha, que esto no es un punto acerca del cognitivismo moral como tal. Más aún, es desde luego válido que podría haber otras consideraciones cognitivistas, de carácter religioso o no, que tuvieran una tendencia más tolerante, pero la perspectiva ampliamente positivista de la cual era parte el subjetivismo de Ayer, estaba dedicada, como una concepción progresista, a rechazar mitos, y ellos se hallaban algunos que institucional o psicológicamente se oponían a la tolerancia. Esto constituye una asociación *histórica*, por lo menos, entre el subjetivismo y la tolerancia.

Dije anteriormente que la teoría prescriptivista y sus similares no encajan fácilmente

en la fenomenología del pensamiento moral. Esto es debido en parte a que las creencias morales que uno tiene no parecen ser cosas que se adquieran por decisión. Más ampliamente, se hace a un lado la intención de tratar de llegar a una conclusión moral. Ayer mismo siempre presenta el asunto moral como algo que sostiene conceptos opuestos a los de otros; el asunto moral no se representa como algo que duda de las concepciones que sostiene. Esto no da suficiente cabida a la parte que se juega en la experiencia moral por la comprensión de que son posibles otras concepciones que podrían ser más satisfactorias; de que otra persona podría estar en lo correcto, y que podría haber explicaciones que demuestren que lo está. En algún grado, estos pensamientos llevan a imágenes cognitivistas u otras imágenes objetivistas: la comprensión de que uno podría estar equivocado y alguien más en lo correcto invita al pensamiento ulterior de que debe haber algo acerca de lo cual podemos estar en lo correcto o equivocados.

Sin embargo, pensamientos de este tipo pueden igualmente motivar a alguien a reflexionar, en una dirección diferente, acerca de las moralidades posibles que no sean como la que uno de hecho tiene, y aquellas reflexiones no favorecen únicamente el objetivismo: pueden conducirnos a algunas asociaciones más amplias entre subjetivismo y tolerancia. “Una moralidad posible” no es aquí sólo un fenómeno humano empíricamente posible, sino algo que podría ser reconocido sobre la base del concepto que cada quien tiene de la vida humana como una solución inteligible a los requerimientos de una sociedad humana en ciertas circunstancias. La idea de que hay varias moralidades posibles es quizá compatible con algunas formas de cognitivismo, pero aquellas formas probablemente representan los “hechos” morales como algo muy general o indeterminado. Sin embargo, si el objetivismo —y no sólo el cognitivismo— es falso, se abre la clase de las moralidades posibles. Esto es opuesto a una suposición de moralidad para todos, identificada por Nietzsche, de que la moralidad es única, la única moralidad; una suposición extrañamente introducida a la estructura del prescriptivismo de Ayer, donde se usa en el modo que ya he mencionado, para acallar el relativismo.

Aceptar que hay varias posibles moralidades, no conduce directamente al relativismo o a la tolerancia. Hay una razón obvia: que el contenido de una moralidad debe determinar hasta cierto punto la actitud de uno hacia otras moralidades en conflicto. Si me opongo a lo que, en mi concepción de las cosas, es llamado el sometimiento de las mujeres, esto no me deja indiferente a los valores de las moralidades que practican, como me parece, el sometimiento de las mujeres. No ser indiferente implica, a su vez, que no seré tolerante. Esto no significa que apoyaré la supresión de aquellas prácticas por la fuerza, pero puedo expresarme contra ellas, exigir restricciones legales, etc. Siendo esto tan obvio, podría bien parecer que la relación entre el subjetivismo y la tolerancia no son más que una asociación histórica. Bien podría ser verdad que el mismo movimiento de las ideas ha apoyado a ambas, y esto es muy comprensible, pero podría aún ser difícil encontrar una conexión más profunda entre ellas. Lo que sucede es simplemente que aquellos que sostienen concepciones subjetivistas también sostienen opiniones morales definidas, algunas de las cuales, en ciertos aspectos, son más tolerantes.

¿Esto es todo lo que puede decirse al respecto? Las conexiones, me parece, pueden llevarse aún más lejos. Primero puede considerarse que, en una explicación subjetivista del asunto, la función de sostener una perspectiva moral es básicamente la de regular y definir las relaciones con las otras personas. En la medida en que esto se entiende con claridad, las perspectivas morales tenderán a perder fuerza si sus expresiones no están dirigidas a personas con las cuales las relaciones de uno necesitan estar reguladas y definidas: en particular, si están dirigidas a personas remotas en el tiempo. El subjetivismo tiende a apoyar lo que en otra parte (1985)[e] he llamado “el relativismo de la distancia”.

Es menos probable que el cognitivismo simpatice con tal perspectiva: si la meta del discurso moral es establecer cómo se hallan las cosas en lo moral, entonces la distancia en sí misma carece de efecto.

Hay algunas consecuencias de esta asimetría, pero no son muy espectaculares. Por un lado, incluso formas consistentes de cognitivismo pueden posiblemente abarcar la idea de que algunas verdades morales son más dignas de ser anunciadas (por un personaje dado en un tiempo dado) que otras; abordar extensamente los méritos de Bruto o los deméritos de Calígula puede parecer tan carente de sentido, si se piensa que se está apuntando a verdades intemporales de la moralidad, como lo es si se piensa que se están haciendo las cosas que Ayer pensó que se estaban haciendo. De hecho, comentarios sobre lo distante pueden cuestionarse más en relación con la práctica general, en relación con perspectivas morales completas, pero entonces es muy improbable que la cuestión sea la *mera* distancia. Habrá importantes diferencias sociales entre las dos situaciones, y entonces el cognitivismo puede desplegar sus habituales recursos para explicar por qué no es necesario adoptar la misma visión para dos situaciones diferentes.

Es verdad que hay una diferencia entre pensar, como podría quizá hacerlo un subjetivista, que las opiniones morales simplemente no se aplican a lo distante, y pensar, como podría quizá pensar un cognitivista, que los juicios a distancia se aplican pero puede no valer la pena anunciarlos. Una dificultad al darle mucha importancia a esto es que los subjetivistas pueden, de hecho, ser reticentes a usar sus recursos distintivos. Como ilustra la propia explicación peculiar de Ayer de nuestra relación con Bruto, una vez que se plantea la cuestión de lo distante, los juicios morales parecen encontrar muy fácilmente la energía para alcanzar el blanco. Un subjetivismo complejo no tendrá dificultad para explicar, a su vez, ese hecho, pero por supuesto no puede al mismo tiempo usar la supuesta asimetría para apoyar su posición: no habrá una asimetría, sino más bien dos explicaciones diferentes de por qué no la hay.

En cualquier caso, un mero relativismo de la distancia no apoyará particularmente la tolerancia. Lo más que podría apoyar sería una forma de indiferencia hacia lo distante. Se puede decir que la indiferencia es, de hecho, inconsistente con la tolerancia: si tú eres indiferente, no necesitas ser tolerante. Pero eso es verdadero, si es que lo es, sólo cuando se asume la tolerancia como una virtud personal, y no la tolerancia como una práctica social y política. La indiferencia es sin duda un camino hacia la tolerancia como una práctica; a duras penas se puede negar que la tolerancia religiosa se ha incrementado con

la declinación del entusiasmo por la religión y los asuntos religiosos. Pero la indiferencia *sólo a lo distante* no promueve la tolerancia. La tolerancia implica esencialmente actitudes hacia quienes no se hallan a la distancia, y los asuntos que plantea son en primera instancia para personas que viven en el mismo espacio social.

Quizá haya una asimetría diferente entre el cognitivismo (al menos) y el subjetivismo, que incide en la situación en la cual la tolerancia puede hacerse necesaria. Puede expresarse en términos de paternalismo. En una concepción cognitivista, si *X* cree en una moral *P*, y *Y* no, *X* puede pensar: si *Y* llegara a creer en *P*, no solamente se tendrían las ventajas resultantes de ese cambio (entre ellas, menos conflicto conmigo), sino que las concepciones de *Y* se encontrarían más cercanas a la verdad. Cuánto contenido hay en este pensamiento depende del tipo de cognitivismo. Pero, si parece haber algún contenido, esto puede (aunque no necesita) proveer a *X* con razones para hacer más probable que *Y* adquiriera esta creencia, en modos quizá ofensivos para el liberalismo. En particular, puesto que *Y* puede considerarse realmente interesado en llegar a saber la verdad, *X* puede recurrir al paternalismo para ayudar a *Y* a que lo logre. Si el cognitivismo es falso, puede no haber razones para justificar el paternalismo. El cognitivismo agrega un posible argumento paternalista para alterar las creencias de otros.

Esto representa una asimetría real que es importante en lo que le concierne. La consideración paternalista adicional no es más que un posible añadido, puesto que el cognitivismo podría tener sus propios argumentos en favor de la tolerancia liberal, como hacer atractiva la capacidad de las personas para encontrar cosas por sí mismas. Esto ilustra un punto general que ya ha aparecido: es improbable que el subjetivismo logre algún resultado importante en esta área. Para cualquier consecuencia del subjetivismo en favor de la tolerancia, habrá alguna versión de cognitivismo u otras formas de objetivismo que arrojen el mismo resultado. (Este es uno de los muchos fenómenos que lo hacen oscuro, conforme este tipo de teorías son progresivamente elaboradas, cuánta distancia va a quedar entre ellas). La única cuestión es averiguar si hay características del subjetivismo que hagan que se derive naturalmente de él, con independencia de la asociación histórica, el apoyo a la tolerancia.

Se podría pensar que este argumento del paternalismo puede ser fuertemente generalizado. Imaginemos una sociedad en la cual hay una variedad de creencias en conflicto sobre cuestiones morales. Entonces, se podría argumentar, el subjetivista, al reflexionar, vería que el intento de hacer que otras partes estén de acuerdo con su propia concepción sería un mero acto de la voluntad, mientras que para el objetivista no lo sería. Ésta es otra aplicación de lo que resulta del argumento del paternalismo; que el objetivista tiene una descripción diferente de lo que ocurre cuando las creencias de los otros cambian. Pero un mero acto de la voluntad de una parte en contra de la otra debe ser inconsistente con las condiciones de coexistencia en sociedad: restringir tales actos de la voluntad es un punto básico de la supervivencia de la sociedad. De ahí que el subjetivismo produzca, se puede sostener a partir de esta obvia premisa, un argumento en favor de la tolerancia del que carece el objetivismo.

Este argumento no puede ser válido como está expuesto. La sociedad está

comprometida a restringir los “meros actos de la voluntad” con procedimientos apoyados por ciertas razones: “lo quiero porque sirve a mis intereses” no se acepta como una razón que se justifique, mientras que “debe ser detenido porque daña los intereses de muchos” puede servir como una razón suficiente, y la acción apoyada por esa razón no será un “mero acto de la voluntad”. Pero cuando el argumento, desde la perspectiva del subjetivismo, plantea la noción de un “mero acto de la voluntad”, esta es una noción que se aplica a cualquier acto que esté basado en una razón moral; o, quizá, a cualquier acto basado en una razón moral que otros no acepten; o al menos a cualquier acto que implique una imposición sobre otro por una razón moral que nadie más aceptaría. La idea de un mero acto de la voluntad que se deriva de una explicación subjetivista de lo que es un juicio moral, es distinta de la idea de un mero acto de la voluntad como lo que está excluido por los requerimientos del arreglo social. Esto no sorprende, puesto que lo primero es, como quien dice, una idea “trascendental”, derivada de un juicio moral, mientras que la segunda está basada en una distinción entre tipos de razones.

¿Se puede mejorar ese argumento? Sólo si el subjetivismo, de un modo u otro, puede contribuir a explicar qué razones podrían, y cuáles no, justificar la supresión o el uso de otro medio de disuasión para hacer frente a comportamientos y creencias morales divergentes. Si preguntamos qué contribución podría hacerse a ese nivel de generalidad, parece que la única respuesta que puede encontrarse es la basada en el requisito de que no se puede ofrecer ninguna justificación que se apoye solamente en los valores de la verdad, o que consista simplemente en la consideración de que es mejor para un grupo creer la verdad moral en vez de la falsedad moral.

Pero esta restricción, más allá de los límites del argumento del paternalismo, puede ejercer muy poca influencia distintiva. En los hechos, después de todo, la afirmación de que en la escuela deberían enseñarse determinadas teorías y no otras, apoyada por el consenso de que dichas teorías son verdaderas, será una afirmación y nada más; habrá apoyo de cualesquiera razones que hagan a uno pensar que aquellas teorías son verdaderas. De modo similar, la afirmación de que algunas creencias morales deberían prevalecer sobre otras —en el límite, que no deberían permitirse las prácticas basadas en las últimas— debe ser apoyada por los objetivistas no con la simple afirmación de que aquellas creencias son verdaderas, sino exponiendo las razones que las justifiquen. En el caso más drástico, la afirmación adicional de que las creencias necesitan ser puestas en vigor tendría que ser apoyada por algo más que la mera afirmación de que son verdaderas. (Si no, volveríamos a estar en el territorio del argumento del paternalismo). Un argumento sobre el asunto será una discusión acerca de, y en términos de, las razones: en primera instancia, será una discusión con las partes forzadas a ceder, y si esa discusión fracasa, entonces tendrán que ser planteadas razones entre las otras partes, las dominantes, para explicar a cualquiera que pueda estar interesado por qué, en este caso, se justifica la coerción.

Todo lo anterior acontecerá sin mayor apoyo de la metaética, ya sea subjetivista u objetivista. Estas actividades, por supuesto, presuponen que la sociedad se halla, a grandes rasgos, interesada en dar razones para justificar sus varias creencias morales y

sus prácticas tolerantes o intolerantes. Todas las sociedades tienen que estar atentas a estos problemas en mayor o menor grado, aunque claramente pueden diferir mucho, en la medida en que es así, en el rango de grupos entre los cuales tienen que explicarse dichas razones, y en el grado de especificidad que es exigido para las políticas adoptadas, en contraste con la atracción general que ofrece la legitimidad. Si una teoría metaética es adecuada para explicar lo que es, desde cualquier punto de vista, tener razones para una actitud moral, parece plausible suponer que tendrá entonces también algunos materiales para analizar estos asuntos. Si una versión del subjetivismo o el objetivismo puede satisfacer el requerimiento más básico de explicar las razones en la discusión moral, entonces también debe ser capaz de dar alguna explicación del papel de dar razones, en particular en una sociedad liberal, y por ende de los requisitos y posibilidades de la tolerancia.

Si este análisis en bosquejo es correcto, parecería entonces que, dentro de la limitada excepción del argumento del paternalismo, pudiera haber una conexión menos intrínseca entre el subjetivismo y la tolerancia de lo que algunos han supuesto. Sin embargo, no se debe olvidar cuán importantes son las conexiones históricas mismas. El mero hecho de que la pregunta pueda ahora presentarse así, y de que algunas versiones adecuadamente elaboradas del subjetivismo y el objetivismo sean ahora difíciles de distinguir, manifiesta que las afirmaciones más fanáticas han sido dejadas de lado. Muy bien podría sostenerse en este punto —y Ayer podría haberlo sostenido— que las simetrías que parecen ahora extenderse sobre más cuestiones como ésta, ya no son simétricas en sus implicaciones, porque son las varias formas de objetivismo, en particular las formas religiosas, las que han perdido su impulso fanático. Puede señalarse que fue tradicionalmente el objetivismo el que se suponía que volvía comúnmente al debate entre objetivismo y subjetivismo, y si ahora parece regresar menos a él, la importancia de ese hecho no es en sí misma simétrica: implica que el objetivismo, en particular, ha perdido algo de su fuerza.

Sin duda, hay algo de paradójico en esta situación, pero el mismo Ayer, ciertamente, hubiera supuesto que lo siguiente es aun más paradójico: si la intolerancia fanática religiosa (y con ella algunas formas particularmente tenaces de objetivismo) no es prominente en la filosofía, eso habla de la filosofía, pero no del mundo. El mismo Ayer hubiera pensado que esto puso a estas discusiones en una perspectiva correcta: dedicó mucha más energía a la defensa de la tolerancia que al análisis de la metaética.

IX. El *actus reus* del doctor Caligari

I

El libro de Michael Moore se subtitula “La filosofía de la acción y sus implicaciones para la ley penal”.^[58] Esta formulación expresa el modo en que procede buena parte de su análisis: una explicación de la acción que se halla filosóficamente (o “metafísicamente”, como él a menudo lo dice) motivada arroja las distinciones y conclusiones que se necesitan para apoyar principios centrales de la ley penal; ese ha sido el caso de la tradición que se remonta a Jeremy Bentham y John Austin. En particular, tres principios fundamentales de la ley penal se defienden del escepticismo filosófico sobre la base de consideraciones filosóficas. Estos son los principios que Moore llama el requisito del acto voluntario, el requisito del *actus reus*, y el requisito del doble riesgo.^[59] Los sutiles e interesantes análisis de Moore tienen éxito al mostrar que estos principios pueden ser defendidos del escepticismo mediante el argumento filosófico.

Sin embargo, hay áreas en relación con las cuales se podría decir que el subtítulo representa una exageración. En estas áreas, lo mejor que puede hacerse es adoptar alguna distinción o conclusión necesaria para la ley penal y mostrar que se puede dar una explicación teórica de la misma que por lo menos no sea inconsistente con formulaciones provenientes de la filosofía de la acción. Parece obvio, al reflexionar, que debe haber áreas en las cuales esto sea verdad. La ley penal, después de todo, tiene metas y propósitos especiales, y es improbable que los requisitos que impone para describir las acciones de las personas coincidan siempre con distinciones que están motivadas muy independientemente de aquellos propósitos especiales. He argumentado en otra parte que cualquier concepción de la responsabilidad implica los cuatro elementos de causa, intención, estado y respuesta.^[60] La responsabilidad es (con ciertas excepciones comprensibles) adscrita uniformemente a una persona como la causa de una situación; surgirán preguntas acerca de lo que la persona estaba tratando de provocar y en qué estado de ánimo se encontraba en ese momento. Finalmente, se hace una asignación con el objeto de responsabilizar a ese agente según las circunstancias.

No hay ningún conjunto de todos estos factores, particularmente los de intención y estado de ánimo, que sirvan a todo propósito, y esto es evidentemente así en la ley contemporánea, donde el conjunto de los requisitos sobre la intención suele ser diferente en la ley de agravio y en la ley penal. Aún más improbable es que la variedad de condiciones para una acción descritas en el lenguaje cotidiano coincidan con los

requeridos por varias ramas de la ley. No hay razón para creer que las varias distinciones que utilizamos habrán de proveer, *ya listas para usarse*, lo que la ley penal necesita, sin tomar en cuenta los requerimientos especiales dentro de la ley penal.

Moore no tiene por qué negar este punto, y frecuentemente parece aceptarlo. Dice, por ejemplo: “la pregunta relevante aquí no es: ¿puede llevarse a cabo cualquier acción compleja ser ejecutada sin la ejecución de un movimiento corporal volitivamente causado? Más bien, la pregunta es: ¿puede cualquiera de las complejas acciones *prohibidas por la ley penal angloamericana* ser ejecutada excepto mediante un movimiento volitivamente causado?”[61] Esto, en sí mismo, no requiere mucho cambio o redirección de la filosofía de la acción en lo que concierne a la ley penal. La pregunta relevante podría encontrar respuesta basándose en fundamentos puramente filosóficos, y la única manera en que la discusión habrá sido conformada a los intereses de la ley penal será restringir la pregunta a una cierta clase de acciones que conciernen a la ley. Sin embargo, incluso este modesto paso requiere la noción del tipo de acciones que son de las que se ocupa la ley penal, y esa noción no va a ser generada por la filosofía de la acción.

Una consideración más importante surge cuando la ley exige respuestas (como, desde luego, frecuentemente lo hace) que los usuarios cotidianos de las descripciones de acción no se sentirían inclinados a dar por razones de sentido común. Consideren las bien conocidas preguntas de dónde y cuándo *A* mató a *B*. [62] Si *A* aprieta el gatillo en una jurisdicción y *B* es alcanzado por la bala en otra, o si *A* administra el veneno en un tiempo y *B* muere como resultado mucho después, hay una notoria dificultad para responder a las preguntas de “dónde” o “cuándo”. La filosofía de la acción —más específicamente, quizá, el análisis de las descripciones de acción— ciertamente muestra por qué hay una dificultad, y, desde luego, que es una condición de si es adecuada tal filosofía el que deba ser capaz de explicar por qué hay una dificultad. Sin embargo, no hay razón para suponer que los procedimientos filosóficos mismos pueden responder a esa dificultad. Esto se debe a que hay una razón especial, como lo es cumplir con los requisitos de la ley penal, para querer contestar la pregunta en primer lugar.

El punto es similar a aquellos planteados por la indeterminación a través de la vaguedad. Es un requisito de la filosofía del lenguaje que pueda explicar claramente por qué es difícil decir cuándo las personas se vuelven calvas o (para poner un ejemplo más contencioso) decir en qué momento un óvulo fertilizado se puede considerar un ser humano. Pero no se debería esperar que aquellas explicaciones respondan a la pregunta. Aristóteles estaba dispuesto a llevar el mismo punto hacia la metafísica misma, cuando dijo que era razonable no buscar más precisión que la que permite el tema subyacente. [63] Pero esta absorción del asunto por la metafísica es, en relación con las preguntas que nos ocupan, ligeramente errónea. La sugerencia de Aristóteles es que si la metafísica de la situación (el tema subyacente) no permite en sí misma que se haga la distinción, entonces ésta no debería hacerse. Pero esto no se deduce —todo lo que se deduce es que la distinción no debería hacerse si uno solamente está interesado en la metafísica. Si ha de hacerse la distinción por alguna otra razón, o cuando, por razones legales, debiera

determinarse dónde o cuándo fue asesinado Smith, entonces uno tiene que ir más allá de la metafísica o la filosofía de la acción a hacer distinciones que uno no puede obtener en aquellos asuntos si se les deja a sí mismos.

II

Moore respeta este punto en gran parte de su práctica. Sin embargo, hay al menos un área en la que trata de hacer que la filosofía de la acción genere una respuesta determinada cuando no lo puede hacer y es explicable que así sea. Hace esto debido a muy obvios requisitos de la ley penal. En esta conexión, quizá podríamos decir que el subtítulo de Moore representa más que una exageración. Ahí, más bien, debería leerse: “la ley penal y sus implicaciones para la filosofía de la acción”. El área en cuestión es la del sonambulismo o del comportamiento bajo hipnosis, y el resto de este artículo estará dedicado a este asunto y a las lecciones que tiene para la metodología general. Este es un asunto en el cual Moore está en desacuerdo con una observación mía (Herbert Hart^[64]) que es apropiada más allá de discusión para describir el comportamiento sonámbulo y el comportamiento similar en el lenguaje, no sólo de la acción, sino de la acción con propósito; he citado a este respecto la maravillosa descripción de Shakespeare de las acciones de Lady Macbeth.^[65] Moore sostiene que “los datos del uso del inglés ordinario”^[66] no nos proporcionan ningún argumento que muestre esto, en absoluto.

Ahora bien, si mi argumento hubiese sido, como implica Moore, simplemente que era “propio del inglés ordinario describir (estos sucesos) con los verbos ordinarios de acción”, no hubiera sido un argumento importante, y Moore se hallaría en lo correcto al llamar nuestra atención sobre el hecho de que los verbos activos pueden aplicarse a sujetos inanimados, como cuando decimos que el ácido sulfúrico disuelve el zinc. Pero no es esto todo lo que había en el argumento. La observación mía era, más bien, que Lady Macbeth había actuado con propósito en varios aspectos y que sus movimientos tenían un contorno intencional.^[67] Por ejemplo, ella se había levantado de la cama, reconoció algo como una luz, la tomó, abrió una puerta, y bajó las escaleras. Todas estas fueron acciones que ella podía haber ejecutado cuando se hallaba despierta, y en buena parte se explican por las razones que se usarían estando ella despierta. Así, abrió la puerta porque tenía el propósito de salir de su recámara. Algo poco usual que hizo fue frotarse las manos como si estuviera lavándolas (cuando no se hallaban en realidad bajo el agua). Esto es algo que Shakespeare capturó de manera exacta:

Doctor: Mire como se frota sus manos...

Caballero: Es una acción que ella acostumbra, para parecer así que se lava las manos.^[68]

Es muy importante para la descripción del caso de Lady Macbeth que, a excepción de la extraña ejecución del lavado de manos, sus acciones son deliberadamente las mismas que las que ella hubiera ejecutado estando despierta, y lo mismo con respecto a las razones que podríamos atribuirle. No es, por lo tanto, simplemente un asunto sobre

un objeto inanimado en una relación causal con su entorno. El problema se plantea precisamente por el hecho de que acciones de este tipo tienen un aspecto intencional o deliberado.

Ahora bien, Moore tiene un modo de reconocer este punto, cuando dice que las acciones de un sonámbulo son “metafóricas”.^[69] ¿Qué es para ellas ser “acciones metafóricas” o, como él también lo formula, “metafóricamente acciones”?^[70] Moore no se está refiriendo a una descripción metafórica de una acción, como cuando decimos que quien objetaba condujo un caballo con todo y coche a través del argumento de quien hablaba. Lo que quiere decir es que los movimientos corporales del agente son descritos metafóricamente como acciones. Él explica esta idea con una teoría peculiar. Atrae nuestra atención a lo que él llama “la frecuentemente no observada ambigüedad en los sustantivos y pronombres por los cuales nos referimos a las personas. Puedo decir: ‘yo golpeo la pelota’, y ‘yo tengo 1.80 de altura’, no obstante, la cosa a la que se refieren estas dos proposiciones es diferente. El primer ‘yo’ se refiere a mí como un agente personal, y el segundo, solamente a mi cuerpo como un objeto inanimado”.^[71]

La explicación de la metáfora, entonces, parece residir en que hay dos objetos de referencia, y podemos usar de un modo metafórico los verbos de acción (que en su sentido literal se aplican a la persona) aplicándolos a un sujeto que es, de hecho, el cuerpo.

Al introducir esta teoría, Moore se refiere en una nota a la distinción de Peter Strawson entre predicados *P* y predicados *M*, ofrecida en el libro de Strawson *Individuals*.^[72] Es más bien irónico que Moore hiciera esta referencia, puesto que la meta de Strawson al usar esta distinción entre predicados fue precisamente evitar una distinción de sujetos —su tesis era que las dos supuestas clases de predicados podían aplicarse a un solo sujeto, o sea, una persona. En cualquier caso, la referencia de la distinción de Strawson no le sirve de mucho a Moore. Nos encontramos con grandes dificultades cuando tratamos de clasificar predicados en dos tipos,^[73] y la idea de Moore implica una versión incluso más severa de la misma dificultad, cuando intentamos determinar dónde en una frase dada ha cambiado la referencia.^[f] “Se sintió apenado, se ruborizó, y empezó a sudar” presumiblemente se refiere a dos objetos de referencia distintos, aunque pudiera ser más bien un poco claro a cuál de ellos se refiere el segundo conjunto. “Se secó con una toalla” no introduce, como podría pensarse, una acción reflexiva, en el sentido estricto de una acción que un agente se haga a sí mismo. Nuevamente, una proposición como “las personas que son muy pesadas tienen dificultades para saltar el muro” requerirá un análisis considerable para descubrir la variedad de cosas de las cuales estamos hablando.

Además de las dificultades de este tipo, se halla el problema básico de que sólo sabremos, en la explicación de Moore, cuál de estos objetos de referencia se está señalando por un pronombre dado cuando entendemos el predicado que se le está aplicando. Hemos sido guiados a la distinción por el uso que Moore hace del ejemplo “yo tengo 1.80 de altura”. Ésta es la pista que seguí al formular las difíciles frases anteriores.

Pero al menos “mide 1.80 de altura” es un tipo de predicado que podría atribuirse a algo que podría no ser una persona —se halla, en la medida en que podamos entender esa distinción, entre los predicados *M* de Strawson. Sin embargo, “acababa de tomar una luz, abrir una puerta, y bajar las escaleras” pareció ser inequívocamente el tipo de predicado que se aplica a una persona. Entonces, si aceptásemos la muy poco convincente sugerencia de Moore de que aparentes referencias a las personas pueden realmente ser referencias a un objeto diferente, o sea un cuerpo, no sería muy probable que tomáramos las proposiciones que indican, al parecer, que las acciones de los sonámbulos son de un corpóreo objeto de referencia. Solamente si hemos *ya* decidido que hay algo peculiar en este predicado, empezaremos a mirar en esa dirección.

La idea de los dos objetos de referencia parece ser evidentemente un artificio inútil. Desde luego, introduce una especie de dualismo que es bastante ajeno a muchas de las cosas que Moore quiere decir acerca de las acciones y sus relaciones con los movimientos corporales. ¿Por qué habrá sido forzado en esa dirección? Más aún, ¿por qué tenía que decir que lo que ejecutan los sonámbulos —o más bien, en su visión dualista, lo que sus cuerpos “inanimados” hacen— son “movimientos corporales involuntarios”[74] que confundimos con acciones? Los movimientos de Lady Macbeth no son, en un sentido fácil de reconocer, movimientos corporales involuntarios. Moore admite que la respuesta al entorno, como el ajuste del movimiento corporal a datos de percepción, se hallan presentes, y estas características, dice, “de hecho hacen que tales movimientos parezcan acciones”. [75] Esto, sostiene, es una ilusión.

Exceptuando perspectivas extremadamente cartesianas, hay algo raro en la discusión de tales casos y su relación con las acciones en términos de *apariencia y realidad*. Un conjunto de movimientos puede, desde luego, verse como una acción de cierto tipo sin ser un ejemplo de una acción de ese tipo; esto es suficientemente familiar en un escenario. Podría haber, desde luego, algún reflejo o movimiento similar que pudiésemos equivocadamente percibir como acción. Pero el caso más obvio en el cual esto podría ser así, sería que no hubiera un conjunto de características con las cuales Moore estuviese de acuerdo, como la conformación del comportamiento con un fin intencional y la sensibilidad a los datos de la percepción; es precisamente la ausencia de tales cosas la que conduciría a uno a decir que no era, después de todo, una acción, sino algo que confundimos con una acción. Pero eso, desde luego, no es el problema con Lady Macbeth. No hay duda de que Lady Macbeth ha tomado la luz, ha encontrado la puerta, la ha abierto, y cuidadosamente ha bajado las escaleras. Más aún, no es un asunto de una rutina mecánicamente determinada que sólo consiste en seguir los datos de la percepción; algunos sonámbulos, al menos —y podemos suponer que entre ellos se halla Lady Macbeth—, caminarán en torno a muebles que no se hallan en su lugar normal. ¿Por qué deberíamos entonces decir que estos movimientos solamente *parecen* acciones?

La razón podría ser, supuestamente, que no tienen un aspecto intencional. Ya he sugerido que esto es incorrecto, puesto que los movimientos que Lady Macbeth hizo al abrir la puerta fueron hechos como movimientos que abren la puerta, y, desde luego, ella ha tenido éxito al abrir la puerta. De modo relacionado, la sonámbula podría tratar de

hacer cosas que no logra hacer: podría tratar de abrir la puerta más de una vez. ¿Hay un argumento, no obstante, para decir que la sonámbula no actúa intencionalmente?

III

Aquí podemos obtener alguna ayuda de los casos de sugestión hipnótica. El hipnotizador le dice al sujeto hipnotizado que haga varias cosas, y muy frecuentemente el sujeto las hace, y puede, desde luego, en los casos más espectaculares, dar varios pasos para hacerlas. Nuevamente, el hipnotizador puede decirle al sujeto que haga una cierta cosa cuando no hay manera de hacer esa cosa, pero el hipnotizador ha hecho que el sujeto crea que la hay. En esos casos, el sujeto puede realizar los movimientos apropiados sin hacer tal cosa. Ahora bien, este sujeto, desde luego, sufre una falsa creencia, y se halla bajo algún tipo de ilusión acerca de su entorno. Pero si ha sido instruido para poner una mesa y, del modo embarazoso que algunos hipnotizadores de escenario imponen a sus sujetos, se le dan objetos muy inapropiados que poner, se puede decir correctamente que no sabe lo que está haciendo.

Ahora bien, podría argumentarse que si se le dice al sujeto hipnotizado que ponga una mesa, y se le dan cuchillos y tenedores reales, aún así no sabe lo que está haciendo, puesto que se halla en un estado en que el hipnotizador podría muy bien haberlo persuadido de que estaba manipulando cubiertos aunque no fuera así. El argumento podría entonces generalizarse para sugerir que los sujetos hipnotizados nunca saben lo que están haciendo. Esto, a su vez, podría sugerir que no hacen nada intencionalmente y, por ende (al menos, sobre principios davidsonianos),^[76] que realmente no hacen nada. Esta es la conclusión a la que quiere llegar Moore, para el hipnotizado y para el sonámbulo.

Éste es probablemente el mejor argumento en favor de la conclusión de Moore, pero no es tan bueno. En primer lugar, hay una paradoja cuando se llega a esta conclusión a partir de estas premisas: la de que ella supuestamente no sabe *qué está haciendo*, y esto desde luego no significa que ella no sabe lo que *su cuerpo* está haciendo. La implicación, más bien, es que ella está haciendo ciertas cosas, pero que no sabe cuáles son.

Esta implicación es seguramente correcta. Incluso, si aceptamos el argumento de que aun cuando ella está manipulando cuchillos y tenedores no sabe lo que está haciendo, y, por ende, no está actuando intencionalmente bajo la descripción *poner una mesa*,^[77] esto es razón para negar que sus movimientos corporales son intencionales; ni para negar que ella hace otras cosas intencionalmente, como moverse hacia un mueble mientras evita otros, ubicando más que dejar caer un objeto sobre una superficie, etc. Desde luego, hay una gran variedad de cosas que el sujeto hipnotizado debe ser capaz de hacer si es que hace lo que sea para obedecer las órdenes del hipnotizador.

El estado de la persona hipnotizada es que actúa bajo las sugerencias del hipnotizador. Un aspecto de esto es que él puede inducir en ella (algo parecido a) creencias falsas. Otro aspecto es que, hasta cierto punto, puede hacer que haga cosas que

ella de otra manera no haría en esas circunstancias. Este segundo aspecto no es simplemente una aplicación del primero: en algunos casos, ella no haría en determinadas circunstancias (en público, por ejemplo, en un teatro) las cosas que le ha hecho pensar que está haciendo. El argumento que hemos considerado descansa sobre el aspecto cognitivo para decir que el agente realmente no actúa, y he negado que esto pueda generalizarse a todo lo que ella hace.

En el caso de Lady Macbeth, no hay ni siquiera un agente que controle, excepto su propio yo culpable, y parte de la razón para describir sus actividades, en términos de propósito e intención, es que son *sus* propósitos e intenciones. Ella comparte algunas fallas cognitivas con el sujeto hipnotizado, y puede muy bien no recordar posteriormente lo que ha estado haciendo. No hay duda de que hay mucho que aprender acerca de estos estados psicológicos, y los descubrimientos de futuras investigaciones pueden desde luego afectar nuestras descripciones. Pero este no es el nivel en el que Moore está operando. Más bien, él supone que podemos ya tener claridad, por razones conceptuales, de que el comportamiento del sonámbulo o el agente hipnotizado no debería ser considerado en términos de acciones. Pero en este nivel precientífico, por el contrario, es claro por qué describimos este comportamiento en términos de acciones, y en tanto que hay diferencias manifiestas con respecto a casos normales de acción, no es difícil simplemente registrarlos.[78] Entramos en dificultades solamente si nos embarcamos en la tarea esencialmente escolástica de insistir en que los criterios de lo que ha de ser una acción son tales que no hay un campo intermedio entre acciones con todas sus características habituales, por un lado, y, por el otro, movimientos corporales “involuntarios”; movimientos que, conforme a la sugerencia de Moore, ni siquiera pueden atribuirse al mismo sujeto como el que en otras circunstancias podría ejercer una acción.

IV

¿Por qué debería alguien querer seguir el curso escolástico? En este caso la razón es obvia; se halla en las exigencias de la ley penal. Incluso si el sonámbulo ejecuta acciones criminales, puede muy bien no ser responsable de ellas o de sus consecuencias. Más aún, podría ser otro el responsable. Podemos trasladarnos del caso de Lady Macbeth —quien, aunque no era inocente, era al menos inocente en su somnambulismo— al de otro caso ficticio: el de la figura criminal del Dr. Caligari, quien en la película *The Cabinet of Dr. Caligari* de Robert Wiene en 1919, controla a otro personaje llamado Cesare (memorablemente actuado por Conrad Veidt).[79] Cesare es mantenido en una caja, y cuando es despertado y guiado por Caligari, se mueve como un sonámbulo a su servicio. Entre otros delitos, mata (como naturalmente lo diríamos) al secretario del ayuntamiento con una daga.

Es interesante que el comportamiento de Cesare es diferente del de Lady Macbeth y del sujeto hipnotizado. A diferencia de Lady Macbeth, su comportamiento se halla bajo

el control de otra persona, pero a diferencia del sujeto hipnotizado, éste no incluye creencias falsas. El doctor Caligari controla sus objetivos, y Cesare hace todo lo que se necesite, a la luz de la situación, para alcanzarlos. Su estado es desde luego peculiar; es un sonámbulo. Pero el argumento, tal y como estaba, para decir que el sujeto hipnotizado no exhibe intenciones o propósitos, tiene menos fuerza sobre él que sobre el sujeto hipnotizado.

Sin embargo, esto no significa que Cesare sea responsable. Supongo que Caligari sería considerado culpable de estas muertes, y no Cesare. Supongo que es bastante obvio por qué la ley puede razonable y justamente llegar a esta conclusión. Aunque no es obvio exactamente qué está mal con Cesare, es obvio que Caligari, en este modo peculiar, controla los objetivos de Cesare. Quiero insistir en que las conclusiones acerca de la responsabilidad no deberían basarse en suponer que la razón por la que Caligari, y no Cesare, es el asesino es simplemente que el asesinato del secretario del ayuntamiento fue una acción de Caligari y no de Cesare.

El secretario del ayuntamiento fue asesinado con una daga, así que, si Caligari lo mató (y nadie más), lo mató con una daga. Quizá lo hizo, y nuestra ligera resistencia a esto puede atribuirse a una implicación cotidiana de decir que Caligari lo mató con una daga, a saber, que Caligari introdujo la daga en la víctima con su propia mano; una implicación que en estas peculiares circunstancias es falsa.[80] Esta implicación, sin embargo, parece ser una *vinculación* de la descripción “lo apuñaló”; así que Caligari no lo apuñaló. Pero si alguien lo apuñaló, entonces (sobre la línea de Moore) fue Caligari, así que tenemos que negar que el secretario del ayuntamiento haya sido apuñalado. Quizá, entonces, deberíamos negar esta implicación. Pero siempre habrá más dificultades del mismo tipo. Cesare ha sido instruido en un repertorio de técnicas para matar, y alguna otra víctima puede haber encontrado su muerte por asfixia, con una almohada mantenida sobre su cara. Caligari debe haberla matado, también. Y mientras, supongo, no la asfixió, más de lo que apuñaló al secretario del ayuntamiento, debe al menos haber provocado su muerte por asfixia, puesto que así fue de hecho como fue asesinada. Pero no lo hizo sosteniendo una almohada sobre su cara; no obstante, seguramente, ella fue asfixiada porque una almohada fue sostenida sobre su cara.

No deberíamos tener que batallar con estas dificultades. Cesare la asfixió, así como apuñaló al secretario del ayuntamiento. Pero lo hizo cuando estaba dormido. Hizo estas cosas como resultado de una muy peculiar relación que tenía con el Dr. Caligari, que le permitía al Dr. Caligari lograr que hiciera aquellas cosas sin que estuviera de acuerdo. Supongamos que Caligari hubiera dicho “¿Estás de acuerdo?”, y Cesare, en su estado sonámbulo, hubiera dicho “Sí, estoy de acuerdo en hacerlo”. Cesare, en realidad, no hubiera estado de acuerdo en hacerlo: *ese* no es un acto que en este estado pueda ejecutar. La explicación ha de encontrarse en su disociación de consideraciones que esencialmente inciden en poder decidir hacer eso. En este estado, él no puede evocar, por ejemplo, pensamientos que relacionen el asesinato con el resto de su vida. Esto ayuda a explicar por qué no es responsable, pero no apoya la idea de que él no haya cometido los asesinatos. No podemos decir que fue forzado a cometerlos; no puso resistencia, y no

tenía conciencia de un precio. No podemos, naturalmente, decir que fue inducido a cometerlos. Podemos mejor decir, quizá, que fue instruido para cometerlos, aunque tendríamos que agregar que recibió las instrucciones en un estado peculiar. Cualquiera que sea la mejor descripción, podemos ver cómo con estos hechos Caligari es culpable de las muertes y Cesare no lo es, pero esto no es porque Cesare no haya apuñalado o asfixiado.

Recordando que se pueden dar “instrucciones” a una computadora, alguien podría decir en defensa de Moore que Cesare podría muy bien haber sido un robot. No está claro exactamente lo que eso significa, pero pienso que no es cierto. Lo que es verdad es que Caligari, para sus propósitos asesinos, muy bien podría haber usado un robot.[81] Hay, de hecho, muchas similitudes entre Cesare y un robot. Pero esto nos dice menos de lo que el defensor de Moore pudiera esperar. No nos dice virtualmente nada, de hecho, a menos que sepamos lo que puede hacer un robot. Lo que es más importante, si hay un descenso de Cesare a un robot (no importa con qué exactitud se describa un robot), hay igualmente un descenso, en la dirección opuesta, de Cesare a Lady Macbeth, a sujetos reales sonámbulos e hipnotizados, a personas que se encuentran por varias razones en estados extremadamente propensos a la sugestión en general o en relación con una persona particular dominante. Debe ser una ilusión escolástica suponer que en algún punto de estos desniveles, cuesta arriba desde Cesare, una acción real, plena, de repente ha de encontrarse en contraposición a un mero movimiento corporal.

Ya tenemos maneras bastante buenas de describir varias dimensiones en las que lo que se hace puede no alcanzar el paradigma de la acción plenamente voluntaria, y la filosofía nos puede ayudar a entender y a desarrollar dichas maneras. Pero debe ser un error suponer que puede proveer por sí misma todo lo que la adjudicación de responsabilidad requiere, en particular la responsabilidad criminal.

He dicho que con respecto a estas muertes, Caligari es culpable y Cesare no lo es. No obstante, Cesare asesinó a ambas víctimas, apuñalando a una y asfixiando a la otra. Así que, ¿de qué es culpable Caligari? Desde luego, de causar o provocar estas muertes. En términos de la estructura cuádruple que he mencionado anteriormente,[82] lo que Cesare hace es una causa de muerte, pero también lo es lo que Caligari hace, que consiste, hemos de suponer, en decirle algo a Cesare. Caligari en el sentido más pleno se propone esa muerte y la tiene como un objetivo. Cesare se la propone sólo en la medida en que la provoca intencionalmente, más que (por ejemplo) por error— aunque es importante que él podría concebiblemente haber hecho eso, en circunstancias que hubieran hecho de ella su error, y no el de Caligari. ¿Mató Caligari a las víctimas? En la propia perspectiva de Moore, Caligari de hecho lo hizo, porque Moore defiende el principio de que “*X* causó la muerte de *Y*” implica “*X* mató a *Y*”. [83] Moore, sin embargo, adopta una perspectiva diferente de la mía con respecto al papel de Cesare, una que le facilita aceptar “Caligari causó sus muertes” y, por ende (sobre su entendimiento de la implicación), “Caligari las mató”. No estoy seguro de, si en mi perspectiva del involucramiento de Cesare —que Cesare las mató—, Moore hubiese dicho que Caligari las mató. Como he dicho, parece difícil negar que Caligari al menos haya causado sus muertes. También pienso que las

mató; mató a la víctima asfixiada, por ejemplo, no asfixiándola, sino creando la situación de que Cesare la asfixiara.

Desde luego, más allá de atribuir la responsabilidad está, que si Caligari es un asesino, es un asesino patológico pero autoconsciente e inteligente, un estado que, desde luego, plantea algunas preguntas acerca de la respuesta que podría ser apropiada para tal persona. Hay preguntas, sin embargo, que surgirían incluso si él no hubiera utilizado un agente sonámbulo. En resumen, las cuestiones de responsabilidad, en el sentido de lo que la respuesta de la ley debería dar a estos acontecimientos, son cuestiones acerca del estado y las intenciones de Caligari y de Cesare, y de la relación entre ellos. No es una cuestión que pueda explicarse atribuyéndole a uno de ellos, y no al otro, acciones e intenciones que fueran causa de muerte.

X. Los valores, las razones y la teoría de la persuasión

1. Razones internas y externas

La distinción que discutiré bajo este título no es, hablando estrictamente, una distinción entre dos tipos de razones, sino más bien una distinción entre dos tipos de aseveración que puede hacerse sobre lo que un agente A tiene razones para hacer. La afirmación “ A tiene razones para X ” se interpreta internamente si se entiende como “ A podría llegar a una decisión de hacer X mediante una deliberación sobre su existente S ”, donde S es el conjunto de deseos, preferencias, evaluaciones y otros estados psicológicos existentes en A en virtud de los cuales puede ser motivado a actuar. Una interpretación externa no conlleva esta implicación. (A lo largo de la discusión adopto la suposición simplificada de que “ A tiene razones para hacer X ” significa “ A tiene razones para hacer X en lugar de otra cosa”; datos adicionales podrían, de hecho, permitirle a uno eliminar esta restricción.) Una opinión que he expresado en otra parte (Williams 1980; 1989; 1995*b*), y que defenderé aquí, es que si “ A tiene razones para hacer X ” tiene un sentido distintivo, entonces debe recibir la interpretación interna; en forma simplificada, las únicas razones para la acción son razones internas. No niego que “ A tiene razones para hacer X ” pueda ser afirmada de manera enteramente inteligible sin esta implicación; simplemente me refiero a que, cuando se afirma así, significa algo que podría expresarse mediante un tipo diferente de proposición, por ejemplo, que es deseable que A hiciera la cosa en cuestión, o que tenemos razones para desear que A la hiciera. Sólo la interpretación interna representa la afirmación de modo distinto como una afirmación acerca de las razones de A . De modo relacionado, si una afirmación de este tipo es verdadera, y A desiste de actuar, lo que es cuestionado es la capacidad de A en este contexto para actuar racional o razonablemente. Sería demasiado fuerte, por más de una razón, decir que en tales circunstancias A actúa irracionalmente, pero la crítica apropiada se dirigirá a la actuación de A como un agente racional o razonable más que otras deficiencias que él pudiera tener.

Es importante enfatizar la variedad de elementos que, desde esta perspectiva, han de incluirse en el S del agente. No sólo consta de inclinaciones o motivaciones egoístas, sino que puede de hecho contener disposiciones asociadas con la percepción del agente de varios tipos de valores. Más aún, no todo lo contenido en el S del agente tiene que estar formado ya en preferencias, y en tanto que está formado en preferencias, éstas no necesariamente tendrían que satisfacer condiciones formales de completitud. No hay

razones naturalistas, basadas en consideraciones de psicología o de filosofía de la mente, para suponer que estas indeterminaciones sean radicalmente reducibles, en particular, a ordenamientos de preferencia que puedan ser manejados por técnicas bayesianas. Si hay una exigencia de tal reducción, será de carácter normativo, más que explicativo, o, quizá, de la fusión de ambos, es decir, una exigencia de que los fenómenos deban hacer posible un particular tipo de explicación.

Esto no es una objeción a la explicación internalista, interpretada con la amplitud que implique vaguedad e indeterminación. Esto sólo refleja la verdad de que proposiciones que afirmen que *A* tiene razones para hacer algo, son en sí mismas vagas y, en diversas formas, indeterminadas. Sería, desde luego, una objeción si la indeterminación fuese tal que no hubiese restricciones a proposiciones de este tipo; si, por ejemplo, no hubiera distinción entre una situación en la que alguien, aconsejando a un agente, lo ayudase a descubrir qué razones tenía para hacer tal cosa, y otra en la que el consejero hiciera una diferencia en el *S* del agente dándole razones que antes no tenía para la acción. Esta cuestión de la vacuidad aparecerá posteriormente en este capítulo.

Hay otra dificultad, que tiene que ver con la forma de la explicación internalista. Al referir a lo que el agente llegaría mediante una sólida deliberación, puede suponerse que si el agente fuera a deliberar, y lo hiciera con solidez, llegaría a esa conclusión. Esto no puede ser totalmente correcto, puesto que la deliberación es en sí misma un proceso psicológico, que tiene efectos.

Esto significa que no podemos necesariamente igualar “*A* podría alcanzar mediante una ruta de deliberación sólida la decisión de hacer *X*” con “*A* llegaría a esa conclusión si deliberara y lo hiciera razonablemente”. Alguien podría, por supuesto, hacer lo que tiene más razones para hacer sin deliberar y, quizá, porque no deliberó. No debemos suponer que pensar qué hacer, revisar consideraciones, etc., sean simplemente percepciones de una realidad externa. Aquí hay un proceso dinámico. Debido a esto, la referencia de la explicación internalista para alcanzar una conclusión mediante una ruta de sólida deliberación tiene que tomarse de una forma idealizada o abstracta, lo que deja fuera efectos psicológicos característicos del proceso de deliberación real. Ha de concebirse como una relación entre el *S* del agente, la acción y el grado de solidez de los principios de deliberación del agente.

Esto plantea la pregunta de qué puede considerarse como principios sólidos, y qué cuenta como una ruta deliberativa sólida. Es esencial para cualquier explicación adecuada de “*A* tiene razones para hacer *X*” que deba ser normativa, y deba admitir la posibilidad de que se le muestre a *A* que está en el error al pensar que tiene razones para actuar en esa forma. La explicación internalista deja espacio para una corrección normativa; permite, por ejemplo, que se le demuestre al agente que su deliberación es inconsistente porque está basada en información falsa. Puede preguntarse por qué la ruta deliberativa del agente puede, desde la perspectiva internalista, mostrarse como inconsistente en referencia a errores de hecho, mientras que las aseveraciones de que lo que el agente está haciendo es inmoral o imprudente, no necesariamente pueden hacerlo. La respuesta a esto es, sin elaborar, que un agente está comprometido, en general, a actuar a la luz de

información razonable, simplemente por ser un agente racional; incluido en el *S* de todo agente racional existe un deseo de no cometer errores. (Hay, desde luego, excepciones complejas a la operación general de este deseo, pero ellas son necesariamente excepciones). Ahora bien, se puede sostener que políticas de prudencia, o incluso morales, se hallan similarmente implicadas en lo que es ser un agente plenamente racional, y algunos filósofos han llegado a conclusiones de este tipo. Puede, desde luego, ser verdadero hasta cierto punto, particularmente con respecto a una cantidad modesta de prudencia; si un agente se halla totalmente carente de una preocupación por la repercusión de sus acciones en sí mismo; podemos, desde luego, tener problemas para entender qué podría contar para él como una ruta de deliberación sólida. Pero en la medida que lo anterior sea verdad, se nos está diciendo algo de los contenidos necesarios del *S* de cualquier agente racional.

Si la aseveración es que aquellos contenidos deben incluir, por ejemplo, elementos morales, necesitamos un argumento para ese efecto; un argumento que los filósofos que han hecho estas aseveraciones han tratado de proveer en algunos casos. Cuando Kant, por ejemplo, sostuvo que un agente racional como tal está comprometido a verse a sí mismo imparcialmente, construyó una filosofía entera con la intención de entregar esa conclusión. Es importante que la explicación internalista en sí misma no excluya una conclusión de ese tipo. Simplemente exige que la conclusión debería argumentarse, y que no puede adquirirse fácilmente, por ejemplo, señalando la obvia verdad de que las personas frecuentemente describen la conducta injusta como “no razonable”. Se requiere algo más que el lenguaje ordinario para entregar grandes conclusiones platónicas, aristotélicas o kantianas en las que la virtud y la razón habrán de coincidir.

2. La racionalidad: interpretación y valor

¿Cuáles son entonces los criterios de racionalidad deliberativa en los que un agente racional podría necesariamente tener interés? ¿Qué tanto trascienden el deseo general de información verdadera y las restricciones mínimas de consistencia? En particular, ¿qué tanto incluyen los tipos de exigencias que la teoría de la decisión racional ha tendido a imponerles? No trataré de responder a esta ambiciosa pregunta, pero hay un aspecto importante de ella que habré de discutir, porque puede parecer que una consideración en esta área socava la perspectiva internalista. Se ha argumentado convincentemente (Hurley 1989; Broome 1991*a*, 1991*b*) que en general no podemos eliminar la ambigüedad de las preferencias de un agente y por lo tanto evaluar, en particular, la coherencia de aquellas preferencias, sin guiar al agente a que tenga valores que deben ser inteligibles para nosotros. Se podría pensar que esto socava la tesis del internalismo de las razones, al requerir que al menos en algún punto “hay una razón para hacer *X*” es prioritario a “*A* tiene razón para hacer *X*”.

El argumento empieza con el punto familiar de que hay contraejemplos *prima-facie*, incluso para principios muy elementales de la teoría de la decisión racional, y que

mientras que es muy fácil salvar la teoría redescubriendo estos casos, se requieren restricciones adicionales si es que este proceso no simplemente ha de volver trivial la teoría. Podemos tomar un ejemplo de Broome (1991*b*: cap. 5) acerca de la transitividad. Maurice, cuando se le da a escoger entre ir a Roma e ir a las montañas, prefiere visitar Roma, porque le asusta el montañismo. Al tener que elegir entre quedarse en casa e ir a Roma, prefiere quedarse en casa, porque es anticultura. De manera compatible con estas dos elecciones, puede preferir ir a las montañas a quedarse en casa.[g] La solución a esta falla de la transitividad se halla en una individuación más fina de las alternativas: podemos distinguir entre la opción de quedarse en casa cuando Roma es la única alternativa y la opción de quedarse en casa cuando el montañismo es la única alternativa. (Sostiene Broome que Maurice prefiere, de hecho, la primera de éstas al montañismo pero esto no es una preferencia práctica, es decir, que se exprese en una elección real. No estoy seguro que esto sea una solución enteramente satisfactoria, y se necesitará alguna redescubrición si es que ha de ser rechazada; pero el principio general de tales métodos es lo que nos concierne aquí). Problemas similares, y soluciones similares, son familiares a partir de la paradoja de Allais y el principio de la cosa segura.

El problema ahora es que tal diferenciación parece ser enteramente arbitraria, con el resultado de que la teoría volverá trivial. La solución de Broome es que Maurice es racional en estas circunstancias si y sólo si se justifica que tenga una preferencia entre las dos opciones subdivididas que se refieren a quedarse en casa; esto implica que no hay un principio racional que le requiere que sea indiferente entre ellas. Esto introduce la noción de una “diferencia justificadora”, que Broome interpreta en términos de una diferenciación con respecto a lo bueno y lo malo. Esto puede parecer que lleva a un rechazo del internalismo. Parece que la alternativa a que no haya restricciones racionales en lo que habrá de contar como una preferencia deliberada, es que las discriminaciones del agente debieran ser racionalmente inteligibles en términos de valores, consideraciones que podemos entender como inteligiblemente relacionadas con cuestiones de valor.

Estoy de acuerdo con Broome en que hay una necesidad de restricciones sustantivas sobre las interpretaciones de la preferencia, pero no estoy de acuerdo en que esto conduzca al rechazo del internalismo. Necesitamos, en primer lugar, distinguir dos posibles metas que pueden atribuirse a un agente racional como tal:

- i*) que llegue a conclusiones acerca de qué hacer mediante procedimientos a los que no se les puede atribuir irracionalidad;
- ii*) que haga lo que tiene mejores razones para hacer.

Podemos afirmar que *i* no es una condición ni necesaria ni suficiente de *ii*. No es una condición suficiente por la sencilla razón de que él puede hacer todo lo que concebiblemente puede, de pedírsele a un agente racional, pero fallar debido a la mala suerte, en particular, debido a ignorancia no culpable, y en ese sentido él no habría hecho lo que realmente tenía las mejores razones para hacer. (Este es el sentido en que “una ruta deliberativa sólida”, si es que ha de proveer una condición suficiente de lo que el

agente tiene la mejor razón para hacer, debe consistir en algo más que en la mejor deliberación posible en ciertas circunstancias). Igualmente, *i* no es una condición necesaria de *ii* por la razón ya mencionada de que la aplicación de cualesquiera métodos deliberativos puede en sí misma, en ciertas circunstancias, tener efectos psicológicos que impiden que el agente haga lo que tiene la mejor razón para hacer. Ahora bien, la meta del agente racional debe ser más básicamente *ii*, en vez de *i*: la idea de que *i* es más fundamentalmente la meta del ejercicio que *ii* es, en efecto, el síndrome notoriamente defensivo del funcionario público.

Lo que nos puede desconcertar sobre esto es que, siguiendo el procedimiento indicado, podremos lograr *ii* solamente porque el problema en cuestión es encontrar qué es la mejor razón de hacer, y que aparentemente esto implica haber logrado *i*. No es un hecho fortuito que *i* ayuda a lograr *ii* pero, aún así, como hemos visto, no siempre se tiene que recurrir a *i* para llegar a *ii*. Ahora bien, supongamos que un agente parece fallar en su búsqueda de *ii* porque alguna relación divergente fue seleccionada para su *S*, por ejemplo debido a una falla de transitividad. (Aquí, desde luego, encontramos el problema habitual de cuál es el fundamento de irracionalidad en tal caso, por ejemplo, si es suficiente decir que se le podría oponer la succión financiera. Es una consideración importante aquí, y que no siempre ha sido reconocida por la teoría económica, que hacer una apuesta es siempre un paso adicional; siempre es una pregunta relevante preguntar “¿Harías un trato de succión financiera con este hombre?”). El interés del agente se halla en alcanzar *ii*: no está pasando un examen en racionalidad práctica. Así que, lo que tiene que hacer es ver si puede explicarse a sí mismo, si puede entender su *S* en términos que racionalicen sus elecciones para sí mismo. Algunas individuaciones finas entre objetos de preferencia sólo están dadas, pero cuando hay un problema real de interpretación, podemos estar de acuerdo con Broome de que ulteriores razones y, desde luego, ulteriores valores, pueden necesitarse en este punto. Pero son los valores del agente, y no del asesor. Son aquellos valores, más aún, a los que un intérprete externo apelaría si es que ha de explicar la conducta del agente como racional.

¿Contradice a Hume esta conclusión? En un sentido, quizá, en tanto que tiene que apelarse a algunos valores para darle sentido a las preferencias,^[84] pero no es antiinternalista, incluso cuando concedamos un punto adicional: que lo que cuente como una invocación inteligible de un valor es una cuestión de interpretación compartida. Hasta cierto punto, puede muy bien ser cierto que podemos interpretar el *S* de otro agente, y, en particular, valores que éste puede invocar al seleccionar sus propios objetos de preferencia, sólo atribuyéndole algunos valores que nosotros mismos compartimos. Esto, desde luego, deja intacta la aseveración internalista de que lo que él tiene razones para hacer es una función de su *S* —sólo dice algo acerca de lo que es entender su *S*.

Sin embargo, la doctrina internalista no tendría sentido si los valores de todos y los *S* de todos fueran los mismos; en ese caso, la deliberación de quien sea podría ser la de cualquier otro, y la distinción entre el externalismo y el internalismo desaparecería, pero claramente no es así. Necesitamos apelar a algunos valores, sin duda, para interpretar las preferencias de otras personas al tratar de entender su racionalidad, pero los valores a los

que tenemos que apelar son los suyos y no los nuestros. Se logrará algo, especialmente si, hasta cierto punto, podemos entender sus valores sólo relacionándolos con los nuestros.

3. ¿Qué tanto excluye la explicación internalista?

Como hemos visto, la interpretación internalista es generosa al definir qué cuenta como una ruta de deliberación sólida. Rechaza la imagen de que un conjunto determinado y fijo de preferencias se exprese simplemente en las ramificaciones racionales de sus decisiones teóricas, e interpreta la deliberación como un proceso de descubrimiento de ellas. La dificultad es que si, al abandonar esta falsa imagen, admitimos que *S* tenga cualquier extensión, perdemos el control de lo que el agente tiene razones para hacer en virtud de dicho *S*. El problema puede modelarse como una deliberación compartida, acerca de lo que *A* ha de hacer, en la que participa *B*, por lo regular viendo las cosas, tanto como sea posible, desde el punto de vista del *S* de *A*. (Algunas de estas consideraciones se aplican también a una deliberación compartida por *A* y *B* acerca de lo que deberían hacer los dos, pero éste no es el caso más relevante para el problema presente). Las limitaciones de lo que cuenta como una expresión de deliberación sólida del *S* del agente se pueden plantear como una distinción entre *A* ayudándole a *B* a alcanzar tal expresión, y *B* intentando cambiar el *S* de *A*. Si abandonamos la idea de que hay[h] un conjunto determinado de procedimientos de deliberación racional (frecuentemente llamado una concepción “algorítmica” de la deliberación), el peligro es, una vez más, que la idea de una “ruta de deliberación sólida” no tendrá ningún efecto moderador, y que cualquier resolución que pueda alcanzar el agente psicológicamente “a partir” de su *S* presente, contará en la misma medida.

La explicación tradicional da por supuesto un *S* muy definido, y un conjunto altamente restringido de procedimientos para alcanzar nuevas decisiones basadas en dicho *S*. En esos términos, uno podría supuestamente definir qué procedimientos de un compañero de deliberación o consejero estarían ayudando al agente a descubrir qué es lo que tiene intenciones de hacer, y no dándole al agente nuevas razones o persuadiéndolo para que haga algo que originalmente no quería hacer. En la medida en que perdamos de vista determinadas restricciones para definir “una ruta de deliberación sólida”, tendremos problemas, también, con estas descripciones derivadas: no quedará claro qué es una asistencia deliberativa para encontrar lo que uno tiene intenciones de hacer, y no una intervención manipuladora.

Una sugerencia posible, en este punto, es invertir el orden de la explicación, y en alguna parte aplicar las restricciones con las que vamos a distinguir los procedimientos que han de contar como asistencia deliberativa y no como intervenciones. Lo que alguien hace porque tiene razones será aquello a lo que puede llegar mediante una ruta de deliberación sólida, y podrá hacerlo, quizá, sólo si él pudiera ser guiado a ella con ayuda deliberativa que opere dentro de aquellas restricciones.

¿Cuáles podrían ser aquellas restricciones? ¿Cómo será el asistente deliberativo útil? Las siguientes pueden ser dos condiciones necesarias:

- i) El asistente será veraz, tanto por decir la verdad como por ayudar al agente a descubrirla. Característico de esto es que el asistente será veraz acerca de sus propios procedimientos y motivos, con el objetivo de que éstos sean transparentes para el agente; el asistente no esconde ninguna intención en sus tratos con el agente.
- ii) El asistente tratará de lograr la máxima comprensión del *S* del agente, y, en particular, si hay un conflicto entre dichas interpretaciones y las del propio agente, el asistente tendrá alguna explicación apropiada de la mala interpretación del agente. (Esta condición no es circular: no implica que el asistente deba lograr la máxima comprensión de cuáles son las razones del agente).

Si un asistente actuando de este modo logra que el agente esté de acuerdo en que lo que tiene mejores razones para hacer es *X*, ¿podemos decir que esto es, seguramente, lo que el agente tiene las mejores razones para hacer, en virtud de su *S* original? No necesariamente. Así como en la construcción puramente individual tenemos que considerar los efectos de las motivaciones psicológicas en la deliberación efectiva del agente, en la situación de dos personas tenemos que tomar en cuenta las motivaciones ajenas a la situación de consejería misma. Si una asistente opera escrupulosamente dentro de las restricciones de las que hemos hablado, y el agente se enamora de ella, radicalmente modifica su *S*. Él es llevado a nuevos cursos de acción como resultado de la deliberación compartida dentro de las restricciones, pero el resultado no es lo que él originalmente tenía razones para hacer.

A pesar de esto, está claro que no podemos descartar totalmente una interpretación retrospectiva de los modos en que una decisión está relacionada con el *S* original del agente. Todavía necesitamos la noción de que la decisión es la expresión, aunque indirecta, de motivaciones que el agente tenía en primer lugar, y no de motivaciones que hubiese adquirido en el proceso, a menos que las motivaciones en sí mismas fueran expresiones de lo que se hallaba ahí antes. No podemos poner todo el peso en el criterio del procedimiento referente a la forma con la cual el asistente ha guiado o podría guiar al agente a una conclusión. ¿Pero hemos puesto ahora algún peso en él? ¿Necesitamos la idea del asistente? Sugiero que sí. Una razón es que el concepto de lo que tal asistente podría hacer juega un papel (aunque solamente un papel) en la interpretación del modo en que la decisión está relacionada con el *S* del agente.

La explicación tradicional de la deliberación racional no da lugar a que dicha deliberación introduzca material radicalmente nuevo. Como uno podría esperar a partir del parecido de la imagen con descripciones comunes de deducción formal, no permite un paso creativo en la deliberación, en el sentido de que el agente piense en un curso de acción que no figure ya en sus creencias o motivaciones. La explicación tradicional podría extenderse de tal forma que podría decirse que, a grandes rasgos, evoca a Popper: podría permitirse nuevo material pero estar interesada solamente en su relación con la

base original, y no con el sitio de donde proviene —consideraría eso como una pregunta puramente psicológica—, pero difícilmente podemos dejar las cosas allí. Una vez que admitimos que la relación entre la decisión y la base original no puede ser algorítmica, o incluso demasiado limitada, se vuelve importante preguntarse cómo el nuevo material entró en la deliberación, y puede pensarse que el hecho de que haya sido favorablemente sugerido por otra persona (por lo menos si no está relacionado de la manera más obvia y directa con la base original) nos da razones para decir que el agente ha recibido una nueva motivación por persuasión y no que haya descubierto lo que ya tenía razones para hacer. Las limitaciones de la naturaleza de la intervención de la otra persona pueden jugar una parte importante en la superación de esta suposición.

Reflexionar sobre el consejo y la ayuda deliberativa nos ayudará a ver no solamente qué es lo que cuenta para ayudar al agente a descubrir lo que tiene razón de hacer, sino también por qué esto no debería, de hecho, compararse con la persuasión.

Los tipos de restricción que se hallan en cuestión aquí no implican que la actividad del consejero no pueda ser una forma de persuasión: puede ser una forma franca de ella, que procure entender del mejor modo los deseos e intereses del agente, y que, por lo tanto, no sea manipuladora. Pero aún así puede llamarse “persuasión”, para recordar que no necesita estar libre de respuestas emocionales; puede funcionar cautivando la imaginación del agente, y puede ser la expresión de un poder utilizado con otros propósitos.

La comparación entre razón y persuasión, tanto en lo individual como en lo político, se remonta, al menos, a Platón, y todavía se halla con nosotros. Tenemos que aprender, considerando lo anterior, que el funcionamiento de la razón, sobre todo en sus aplicaciones prácticas, no puede reconocerse sobre bases formales; e igualmente, que una explicación del procedimiento que intente eliminar la dimensión de la persuasión tendrá que estar vacía (este punto se aplica, obviamente, a la “situación del discurso ideal” de Habermas, una cuestión que es muy relevante para el tema presente pero que no puedo abordar aquí). Aun cuando las razones en cuestión son las de un individuo y, por lo tanto, como ya lo he dicho, están basadas en su *S*, podemos ver por qué están conectadas las dos ideas; la del agente que descubre lo que tiene razones para hacer y la de un asesor comprensivo que opera dentro de un particular tipo de restricción. Si seguimos tales conexiones, podremos ver más claramente por qué incluso las razones individuales internas requieren una discusión con otras personas; por qué también, aunque esperamos que la discusión política pública sea moderadamente racional y aborde las razones de los individuos y los grupos, no tenemos por qué suponer absurdamente que podría trascender el mundo de la persuasión y del poder. Desde luego, la explicación interna de las razones del individuo nos muestra particularmente la importancia de vincular el análisis de las razones con la teoría de la persuasión: las fallas de una explicación externalista aparecen de la manera más clara cuando reflexionamos sobre el tipo de argumento que podría necesitarse para convencer a un agente de que ciertas razones son suyas, y cómo es que eso difícilmente podría dejar de ser manipulador.

Referencias

- Allais, M., “The so-called Allais Paradox and Rational Decision under Uncertainty”, en M. Allais y O. Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*, Reidel, Dordrecht, 1979.
- , y O. Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*, Reidel, Dordrecht, 1979.
- Broome, J., “Rationality and the Sure-Thing Principle”, en G. Meeks (ed.), *Thoughtful Economic Man*, Cambridge University Press, Cambridge, 1991a.
- , *Weighing Goods*, Basil Blackwell, Oxford, 1991b.
- Harrison, R. (ed.), *Rational Action*, Cambridge University Press, Cambridge, 1980.
- Hurley, S. I., (1989), *Natural Reasons* (Oxford: Oxford University Press).
- Meeks, G. (ed.), *Thoughtful Economic Man*, Cambridge University Press, Cambridge, 1991.
- Williams, B., “Internal and External Reasons”, en R. Harrison (ed.), *Rational Action*, Cambridge University Press, Cambridge, 1980, reimpresso en B. Williams, *Moral Luck*, Cambridge University Press, Cambridge, 1991.
- , “Internal Reasons and the Obscurity of Blame”, *Logos* 10, Santa Clara, California, 1989, reimpresso en B. Williams, *Making Sense of Humanity*, Cambridge University Press, Cambridge, 1995.
- , *Making Sense of Humanity*, Cambridge University Press, Cambridge, 1995a.
- , “Replies”, en J. E. J. Altham y R. Harrison (eds.), *World, Mind, and Morality: Essays on the Ethical Philosophy of Bernard Williams*, Cambridge University Press, Cambridge, 1995b.

XI. Responsabilidad moral y libertad política[i]

En el concepto “responsabilidad moral”, el término “moral” puede tener dos significados distintos. Por un lado, puede referirse a un campo particular de aplicación y a sus correspondientes consecuencias, que son informales y sociales más que formales y apoyadas por la fuerza. En esta acepción, “responsabilidad moral” se distingue de responsabilidad legal. En un uso muy diferente del término, “moral” implica cierta base de evaluación que pone particular énfasis en lo voluntario. En este sentido, la responsabilidad moral puede ser propuesta como una base de evaluación, incluso cuando lo que se busca es la responsabilidad legal, por ejemplo, en relación con la ley penal.

Cuando la primera idea es dominante, y se enfatiza el contraste con (en particular) lo legal, no tiene que enfocarse especialmente en la voluntariedad. Se podría decir de alguien que es moral, pero no legalmente, responsable por una omisión, o por las consecuencias de haber llevado irresponsablemente a alguien a formarse una expectativa particular. Sin embargo, cuando se exige en el segundo sentido, que la responsabilidad moral sea la base para evaluar la responsabilidad en la ley penal, se enfatiza la voluntariedad. Éste será el tema que nos ocupe en el presente artículo.

He argumentado en otra parte[85] que cualquier concepción de la responsabilidad incluye cuatro elementos: causa, intención, estado y respuesta.

Si la relación de un agente con un resultado dado ha de concebirse como *responsabilidad*, es decir, si no se ve a dicho agente como chivo expiatorio, éste debe, básicamente, considerarse como la *causa* del resultado. Hay excepciones a eso, cuando mediante maniobras legales la responsabilidad se transfiere, atribuye, supone, etc., de manera que puede ser eliminada, como en la mayoría de las situaciones de inculpación restringida. La presente discusión está confinada al caso básico, la responsabilidad personal en la que el agente es la causa del resultado en cuestión.

La distinción entre lo que un agente pretende hacer (si es que pretende hacer algo) y el estado mental en que se encuentra cuando lo intenta y lo hace, es la diferencia entre *intención* y *estado*. [86] La distinción es importante, aunque en casos particulares puede ser dudarse si estamos enfrentándonos con la intención, el estado o ambos. Un hombre que en un ataque de furia destruye un documento vital e irremplazable puede, en ese momento, no haberse dado cuenta de lo que estaba haciendo, o puede haber sabido lo que estaba haciendo y en ese momento no haberle importado, o puede no haber razones para decir una cosa o la otra. En ciertos estados, la acción intencional es imposible. Michael Moore[87] así lo cree para todos los estados de sonambulismo e hipnóticos, pero

es seguramente una exageración; si es importante para la ley penal saber si un agente se encontraba en uno de esos estados, no es porque éste no haya podido hacer nada intencionalmente, sino más bien porque los actos intencionales que habría hecho en uno de esos estados no tienen una relación normal con sus planes o su carácter.

Respuesta significa lo que es esperado, exigido o requerido del agente, o lo que se le impone. En algunos casos, particularmente en la ley penal, eso se llama castigo. El presente análisis se concentrará en este ejemplo, pero es importante aclarar que no es la única respuesta reconocida en relación con la responsabilidad ni, desde luego, con la responsabilidad moral.

Una concepción de la responsabilidad penal como responsabilidad moral requiere que el agente tenga ciertas *intenciones* y se encontrara en cierto estado que justifique que se diga que sus actos son voluntarios. “*A* hace *X* voluntariamente” es equivalente a “*A* hace *X* intencionalmente, en un estado mental normal”. La inherente vaguedad de esto nos ocupará posteriormente.

Es una pregunta importante y bien conocida, en relación con el divorcio voluntario, los seguros de automóviles, etc., la de si en una área dada tiene sentido considerar el concepto de responsabilidad. Sin embargo, aquellos que ponen énfasis en la noción de responsabilidad moral no suelen ver las cosas de este modo. Se piensa que la responsabilidad moral se aplica, como quien dice, a sí misma. La consecuencia de esta actitud es que se piensa que si la causalidad, la intención y el estado son de un cierto tipo —si el agente es voluntariamente la causa de ciertos tipos de resultado—, entonces en sí mismo atrae la respuesta apropiada. Donde la respuesta es el castigo, ésta es una versión[88] de la teoría retributiva del castigo. Este “internalismo” o autosuficiencia de criterio es una característica general de las concepciones asociadas con la moral. Aquí, como con el juicio moral y el mérito moral, el sistema de moralidad intenta sostener la idea de que no hay un punto de vista externo relevante desde el cual puedan ser evaluadas sus características.

Esta idea no puede ser sostenida enteramente; necesitamos una perspectiva externa razonable de las prácticas y disposiciones de la responsabilidad moral. Esta última obtiene algún apoyo del hecho de que la mejor alternativa después de ella en la ley penal es un tratamiento utilitarista directo de los casos particulares, y esto es mucho menos atractivo que la responsabilidad moral. Sin embargo, como señalara Hart,[89] los argumentos consecuencialistas se pueden aplicar a un nivel más general. El argumento utilitarista de alto nivel presentado por Bentham, para restringir el castigo a las ofensas voluntarias —porque solamente las ofensas voluntarias pueden impedirse— no es razonable, puesto que la práctica de castigar lo involuntario puede, de hecho, ayudar a impedir las ofensas voluntarias, a través del miedo. Sin embargo, Hart ofreció un mejor argumento, que invoca no solamente el bienestar sino también la libertad. Éste está basado en el principio de que los ciudadanos deberían ser capaces de conducir sus asuntos en la medida de lo posible sin que el poder del Estado se dirigiera impredeciblemente contra ellos. Este argumento considera también otras cosas, como evitar la legislación retrospectiva y lograr que las leyes sean comprensibles y, en la medida de lo posible, bien conocidas.

Este argumento se relaciona muy directamente con el valor de la libertad *política*: es decir, la libertad del ciudadano con respecto al uso intencionalmente dirigido del poder estatal. No se refiere simplemente a maximizar la libertad del individuo con respecto a la presión social. Uno de los efectos del delito no controlado es disminuir la libertad de las personas, y un cálculo maximizador directo tendría que sopesar estos efectos en contra de las limitaciones a la libertad impuestas por la disposición de la ley a castigar lo no voluntario.[90] El argumento implica, más bien, una responsabilidad propia del Estado con respecto al uso del poder; una responsabilidad que, supongo, será reclamada por cualquier versión del liberalismo. Éste no es el lugar para exponer una teoría política del liberalismo; el punto es nada más que este aspecto de la responsabilidad moral pertenece a la teoría política.

El argumento expresa específicamente una preocupación por el poder estatal, y no por cualquier otro poder que pudiera limitar la libertad de los ciudadanos, pero se halla también, muy en especial, preocupado por el uso intencionalmente dirigido del poder estatal. Se puede decir que los ciudadanos de cualquier manera sufren en un grado indefinido los efectos colaterales del poder estatal. Esto es verdad, pero hay que distinguir, con respecto a la libertad, entre ser objeto del poder de alguien y ser afectado incidentalmente por sus actividades como se desprende de las muy primitivas conexiones entre carecer de libertad, hallarse bajo el poder de otra persona y estar sujeto a la voluntad de esa persona. La condición de ser políticamente libre, en cierto sentido, no es la de no ser afectado por lo que hace el Estado, sino más bien la de que las intenciones del Estado no se conviertan forzosamente en las de uno. Nuevamente, por qué esta distinción es importante, es algo que concierne a la teoría política.

Entonces es un argumento de la filosofía política el que proporcionará consideraciones externas que apoyen los criterios de responsabilidad moral. Este argumento se aplica especialmente al castigo penal, con respecto al cual no puede haber algo como un seguro contra la imputabilidad. No se puede tramitar un seguro para compensar a las víctimas de actos criminales y, aunque el Estado pudiera hacerlo, tendría muchos límites; también hay un límite en la justicia política, puesto que es probable que represente una pérdida común de seguridad en beneficio de la voluntad del delincuente, si no es que del delincuente mismo.

El argumento a favor de la responsabilidad moral en relación con el castigo penal, es el siguiente: concedida una justificación del castigo tal como la disuasión, las exigencias de transparencia implican un sistema de castigos públicos aplicados a los delincuentes; la libertad política implica, más aún, que debería ser aplicada sólo a los que han actuado voluntariamente.

Está claro, sin embargo, que el castigo no puede explicarse enteramente por la disuasión. Mucho en él es una expresión de enojo, y cuando se conduce propiamente, es considerado como la justa expresión del enojo justo. Esta verdad podría conducir a una “teoría retributiva” del castigo. Está menos claro que nos lleve a una justificación retributiva del castigo. Más bien, podría ser que si uno aceptara la teoría retributiva, dejaría de buscar una justificación del castigo como tal, aunque pueden muy bien

buscarse justificaciones de prácticas específicas: uno puede, por ejemplo, oponerse a largos intervalos entre acto y castigo. Una dificultad sería es que puede resultar que ningún procedimiento justo en un Estado moderno pueda adecuadamente expresar la retribución, y éste es un punto que nos ocupará posteriormente.

Las llamadas “justificaciones retributivas” del castigo frecuentemente parecen descripciones didácticas de lo que se hace al castigar por retribución, como cuando se dice “eso le enseñará” o “le dará una lección”.[91] Éstas son descripciones consecuencialistas erróneas de lo que es, de hecho, en la explicación retributiva, un proceso intransitivo. Basándonos en la comprensión de la retribución, no se *hace* esencialmente nada al castigar mas que lastimar al receptor. Se sostiene que es lastimado con justicia, pero ese criterio no es lo que justifica el castigo, sino lo que él ha hecho.

Vale la pena decir que una teoría puramente denunciatoria del castigo no parece ser una teoría del castigo, a menos que la denuncia sea en sí misma lo suficientemente dolorosa como para ser un castigo. La idea de que los castigos tradicionales dolorosos son simplemente denuncias, es incoherente, porque no explica, sin caer en petición de principio, por qué las denuncias tienen que asumir la forma de lo que Nietzsche identificaba como la constante del castigo, “la ceremonia del dolor”.

Con un entendimiento retributivo del castigo, aún procede el argumento a favor de la responsabilidad moral, fundada en la libertad política. (Hay una diferencia en el argumento cuando se expresa más como retribución que como disuasión. Cuando la justificación es la disuasión, necesitamos una premisa extra acerca de la transparencia, puesto que los mecanismos sociales como la disuasión, como los utilitaristas frecuentemente nos recuerdan, no son forzosamente efectivos; la retribución legal, por otra parte, es esencialmente pública).

El argumento de la libertad política como criterio de la responsabilidad moral se halla todavía disponible bajo un entendimiento retributivista del castigo; lo que es más importante, todavía es necesario. Esto no se debe a que necesitemos todavía un argumento a favor del castigo. Para una perspectiva retributivista, no se necesita argumentar ya el castigo (en principio) que va a imponerse —aunque permanece la pregunta importante de si los procedimientos de defensa legal también pueden expresar adecuadamente la retribución. La razón por la que todavía necesitamos un argumento a favor de la responsabilidad moral es que la retribución no necesita forzosamente estar dirigida a actos voluntarios. Esto se pasa fácilmente por alto, porque la presencia de las intenciones del agente son importantes para la psicología de la retribución. No obstante, el carácter voluntario no puede ser estrictamente necesario, puesto que los sentimientos retributivos fácilmente se extienden a un agente imprudente por el resultado real de sus actos, y no simplemente por su imprudencia.

Más aún, incluso cuando el agente tiene la intención de hacer daño, la retribución es menos sensible que la responsabilidad moral a su *estado* psicológico. Un aspecto de la “atrocidad” es el horror del acto, y con acciones horribles las consideraciones retributivas pueden ser bastante insensitivas para el Estado. Sin embargo, otro aspecto de la “atrocidad” es la premeditación, y algunas de las consideraciones asociadas con ella están

en los límites entre estado e intención, como en los actos impulsivos: ¿el estado del agente era tal que no sabía lo que estaba haciendo, o bien, aun sabiéndolo momentáneamente no le importó?

Si pensamos que el castigo retributivo debería ser restringido por atención a la responsabilidad moral, esto no se deduce simplemente de lo que es la retribución (al menos por lo que concierne a muchas preguntas acerca del estado del agente), sino más bien del argumento ya utilizado en relación con la disuasión, el argumento que se basa en la libertad política.

Como ya he mencionado, hay una cuestión muy seria acerca de la operación del castigo retributivo en una sociedad moderna. Una verdad parcial contenida en explicaciones expresivas o denunciatorias del castigo es que la retribución tiene que ser vista como tal si es que ha de tener algún caso, pero si el proceso judicial requiere dar oportunidades de apelación que duran mucho tiempo, y si, en relación con esto, se imponen limitaciones a las formas que puede asumir el castigo, no queda claro que lo que realmente sucede tenga que ver con la idea de retribución de cualquiera. En un sentido, el paradigma del castigo retributivo es un linchamiento, bajo la condición de que sea el culpable a quien se ahorca. Notoriamente, nunca fue fácil que se cumpliera esa condición. Bajo las ideas modernas de lo que será asegurarse de satisfacer ésa y otras condiciones, incluso en una ejecución, si es que se permiten las ejecuciones, no es posible alcanzar el estándar expresivo de un linchamiento. Pero no hay marcha atrás en las exigencias modernas sobre la justicia en sus líneas generales, incluso si se pueden modificar prácticas particulares. El punto aquí no es que la retribución sea inherentemente mala, primitiva o irracional; es simplemente que ningún castigo bajo las condiciones modernas puede realmente ser una retribución. Aquellos que piensan que el castigo tiene mucho que ver con la retribución deberían, en consecuencia, preguntarse si cualquier cosa que se le haga a un ofensor bajo las condiciones modernas puede ser realmente un castigo.

Incluso si tenemos un argumento general a favor de la responsabilidad moral, debemos reconocer que sus efectos se hallan severamente restringidos por las limitaciones del concepto mismo de lo voluntario. Ese concepto es efectivo hasta cierto punto, pero es tan vago como superficial. Éstas son dos consideraciones diferentes. Hay una gran falta de determinación y vaguedad acerca del estado psicológico del agente y también sobre la solidez de los atenuantes relacionados, como en el caso del “impulso irresistible” y otras áreas de oscuridad que se hallan entre estado e intención. Hay preguntas que no se pueden responder acerca de lo que es “normal” para un agente dado, y disputas acerca del grado en el que eso debe tomarse en consideración.

El concepto de lo voluntario es también superficial, porque incluso si se establece que la acción del agente era intencional y que él estaba en un estado normal (*e.g.* altamente premeditado), hay una cuestión ulterior: por qué puede querer hacer tales cosas, si se sabe a ciencia cierta que él es tal persona, etc. Es claro que si la voluntariedad ha de funcionar, tales cuestiones no pueden llevarse más allá de cierto punto. No es que obtengan una respuesta favorable, como suponen los partidarios absolutos del libre

albedrío, ni que obtengan una respuesta desfavorable que ponga la responsabilidad moral fuera de discusión. Ni tampoco es, como algunos reconciliadores quizá suponen, que haya un motivo evidente para interpretar lo voluntario dentro de ciertos límites que excluyen aquellas preguntas. Es simplemente que lo voluntario es un concepto inherentemente superficial, al cual no se le debiera exigir demasiado.

No supongo que este hecho haga al castigo, bajo esta concepción poco profunda de la voluntariedad, injusto en sí mismo; pero hay una cuestión real acerca de lo que puede significar el castigo, y de si puede significar lo que la mayor parte de las personas quieren que signifique. Con el castigo, como Nozick correctamente lo dijo, pasa lo mismo que con el premio: no llega hasta donde debe llegar.[92] Podemos extraer algunas lecciones de esto para la responsabilidad moral. Podemos usar la noción, pero no debemos llevar las cuestiones asociadas con ella demasiado lejos; no debemos insistir en sus méritos en los casos en que éstas son puestas a prueba, si hay algo más que satisfaga de mejor manera nuestras necesidades, y por sobre todo, no debemos crear un misterio excesivo con él.

Por supuesto, en esa última recomendación hay un juicio significativo. Es perfectamente posible que las personas estén de acuerdo con la mayor parte de lo que he dicho, pero estén en desacuerdo con este juicio, porque piensan que el único modo de mantener la responsabilidad moral es convertirla en un misterio excesivo. No creo esto, porque no pienso que en el mundo moderno se pueda actuar con impunidad, y porque considero que los valores, que hacen imposible actuar con impunidad también penalicen el querer actuar con impunidad. Tanto yo como mi oponente hipotético rechazamos los intentos tradicionales, en especial los de la ilustración, de racionalizar completamente la responsabilidad moral, pero la actitud que se tiene hacia la ilustración, aquí como en cualquier otra parte, es finalmente un asunto de la actitud que se tiene hacia la honestidad política.

La responsabilidad moral tiene una función, y hay muchas razones para decir que ayuda a imputar culpas y a encontrar el modo de responder a las ofensas, en tanto que uno no la tome demasiado en serio. Pero es muy frecuente pensar que la responsabilidad moral es autoaplicativa y no necesita ni permite ninguna justificación o evaluación externa; que es profunda o puede hacerse así; que es el ideal en términos del cual han de ser juzgadas las demás instituciones y a la cual debieran tratar de aproximarse. Todas estas concepciones adicionales son falsas. Si la responsabilidad moral necesariamente implicara estas concepciones, tendría que ser abandonada.

XII. Tolerar lo intolerable

La dificultad con la tolerancia es que parece ser al mismo tiempo necesaria e imposible. Es necesaria donde diferentes grupos tienen creencias en conflicto —morales, políticas o religiosas— y se dan cuenta de que no tienen más alternativa que vivir juntos; es decir, excepto el conflicto armado, el cual no resolverá sus desacuerdos e impondrá un sufrimiento continuo. Éstas son las circunstancias en las que la tolerancia es necesaria. No obstante, en esas mismas circunstancias podría, muy bien, parecer imposible.

Si la violencia y el derrumbe de la cooperación social son amenazados en estas circunstancias, es porque las personas encuentran las creencias o modos de vida de los otros profundamente inaceptables. En asuntos de religión, por ejemplo (que históricamente fue la primera área en que se usó la idea de tolerancia), surge la necesidad de la tolerancia porque uno de los grupos, al menos, piensa que el otro está equivocado de un modo blasfemo, desastroso y obscuro. Los miembros de un grupo pueden también pensar, muy frecuentemente, que los líderes o ancianos del otro grupo impiden la ilustración y la liberación de los jóvenes o quizá de las mujeres. En este caso, ven que no es sólo en el interés de su propio grupo, sino también en el de algunos miembros de otro grupo, que debería prevalecer la verdadera religión (la que ellos creen que lo es). Es porque el desacuerdo se vuelve tan profundo que las partes enfrentadas piensan que no pueden aceptar la existencia del otro. Necesitamos tolerar a otras personas y sus modos de vida solamente en situaciones que hacen muy difícil hacerlo. La tolerancia, podemos decir, se requiere solamente para lo intolerable. Éste es su problema básico.

Podemos pensar en la tolerancia como una actitud que un grupo más poderoso, o una mayoría, puede tener (o puede dejar de tener) hacia un grupo o una minoría menos poderosa. En un país donde hay muchos cristianos y pocos musulmanes, puede surgir el problema de si los cristianos toleran a los musulmanes; los musulmanes no pueden elegir, por así decirlo, tolerar o no a los cristianos. Si las proporciones de cristianos y musulmanes se invierten, así lo hará también la dirección de la tolerancia. Así es como solemos pensar en la tolerancia, y es natural hacerlo, porque las discusiones sobre la tolerancia frecuentemente han versado acerca de las leyes que deberían existir —en particular, que permitan o prohíban varios tipos de práctica religiosa—, y las leyes han sido determinadas por las actitudes del grupo más poderoso. Pero más básicamente, la tolerancia es cuestión de las actitudes de cualquier grupo hacia otro, y no tiene que ver solamente con las relaciones del más poderoso con el menos poderoso. De hecho, no se trata simplemente de definir qué leyes debería haber. Se podría decir correctamente que

un grupo o credo sería “intolerante” si quisiera suprimir o expulsar a los otros, incluso si, de hecho, no tuviera el poder para hacerlo. Los problemas de la tolerancia han de encontrarse primeramente en el nivel de las relaciones humanas y de la actitud de un modo de vida hacia otro. No es solamente cuestión del modo en que debe usarse el poder del Estado, aunque, desde luego, tiene que ver con eso —como problema de filosofía política—. Sin embargo, no hay que apresurarnos a concluir que lo que subyace a la *práctica* de la tolerancia debe de ser la *virtud* personal de tolerancia. Toda tolerancia implica dificultades serias, pero es la virtud que más drásticamente amenaza con implicar una imposibilidad conceptual.

Si ha de haber un problema de tolerancia, es necesario que haya *algo que tenga que ser tolerado*; tiene que haber alguna creencia, práctica o modo de vida que un grupo pueda pensar (no importa cuán fanática o irrazonablemente) que es equivocado, erróneo o indeseable. Si un grupo simplemente odia a otro, como ocurre con una vendetta de clanes, o en casos de puro racismo, lo que se requiere no es solamente tolerancia: las personas implicadas necesitan más bien superar su odio, sus prejuicios o sus recuerdos terribles. Si estamos pidiendo que las personas sean tolerantes, estamos pidiendo algo más complicado que eso. Desde luego, deberían dejar de lado su deseo de suprimir o de expulsar a la creencia rival, pero también mantendrán su compromiso con sus propias creencias, que es lo que les dio ese deseo en primer lugar. Hay una tensión aquí entre los compromisos de uno y la aceptación de que otras personas pueden tener otros compromisos quizá muy desagradables. Ésta es la tensión típica de la tolerancia, y la tensión que la hace tan difícil.

Precisamente porque tiene que ver con esta tensión entre el compromiso con la perspectiva propia y la aceptación de la del otro, se supone que la tolerancia es algo más que mero hastío o indiferencia. Después de las guerras europeas de religión en los siglos XVI y XVII, las personas empezaron a pensar que debía ser mejor la coexistencia para las diferentes iglesias cristianas. Con ese desarrollo vinieron varias actitudes nuevas. Algunas personas se volvieron escépticas acerca de las aseveraciones distintivas de cualquier iglesia, y empezaron a pensar que no había verdad, o al menos que no había una verdad descubrible por los seres humanos, acerca de la validez del credo de una iglesia en oposición al de otra. Otras personas empezaron a pensar que las luchas les habían ayudado a entender mejor los propósitos de Dios: que a él no le importaba de qué manera la gente rendía culto, en tanto que lo hicieran de buena fe dentro de ciertos límites cristianos amplios. Y en tiempos más recientes, un espíritu ecuménico similar se ha extendido más allá de los límites de la cristiandad.

Estas dos líneas de pensamiento fueron en cierto sentido en direcciones opuestas. Una de ellas, la escéptica, penso que podía conocerse menos acerca de los designios de Dios de lo que habían supuesto los partidos en guerra, cada uno con su particular fanatismo. La otra línea de pensamiento, la perspectiva de la Iglesia amplia, pretendió tener una mejor compenetración en los designios de Dios que habían tenido los partidos en guerra, pero en su relación con las batallas de la fe, las dos líneas de pensamiento terminaron, no obstante, en la misma posición: que las cuestiones precisas de la creencia

cristiana no importaban tanto como las personas habían supuesto; que había menos juego. Esto condujo a la tolerancia como una cuestión de *práctica* política, y ése es un resultado extremadamente importante. No obstante, como actitud, es menos que tolerancia. Si no importa tanto lo que cualquiera cree, no se necesita la actitud de la tolerancia más de lo que se necesita con respecto a los gustos de comida de otras personas.

En muchos casos, las actitudes que son más tolerantes en la práctica surgen por esa razón, que las personas dejan de desaprobar o de tener un juicio negativo hacia un cierto tipo de comportamiento. Eso es lo que está sucediendo, en muchas partes del mundo, con respecto a tipos de comportamiento sexual que previamente fueron desalentados y, en algunos casos, legalmente castigados. Una relación extramarital o un *ménage* homosexual puede no hacer surgir ningún comentario o reacción hostil, como sucedía en el pasado, pero una vez más, aunque esto es tolerancia en la práctica, la actitud en la que se basa es la indiferencia más que, estrictamente hablando, la tolerancia. Desde luego, si yo y otros en el vecindario dijéramos que estamos *tolerando* las relaciones homosexuales de la pareja que vive en la casa de al lado, se pensaría que nuestra actitud es bastante menos que liberal.

Hay, sin duda, muchos conflictos y áreas de intolerancia de los cuales, desde luego, la mejor solución es el incremento de la indiferencia. Comportamientos sexuales y sociales que en sociedades más pequeñas y tradicionales son de gran preocupación pública vienen a parecer ahora más un asunto privado, que no hace que nadie se pregunte si son correctos o incorrectos. Caer en la indiferencia puede también proporcionar, como sucedió en Europa, la única solución a algunas disputas religiosas. No todas las religiones, por supuesto, buscan convertir a otros, ya no digamos ejercer coerción sobre ellos. Sin duda tienen alguna opinión u otra (quizá del tipo de la “Iglesia amplia”) acerca del estado de verdad o error de quienes no comparten su fe, pero se contentan con no molestarlos. Otros credos, sin embargo, están menos dispuestos a permitir que lo que consideran error florezca, y puede ser que con ellos no haya solución excepto la que Europa descubrió (en religión, al menos, si no en política): una disminución del entusiasmo. Es importante que dicha disminución no se vea como un mecanismo al que se le agota el combustible. Como descubrieron las diversas sectas de la cristiandad, una religión puede tener sus propios recursos para repensar sus relaciones con otras. Una idea relevante, que tuvo considerable influencia en Europa, es que una religión expansiva realmente quiere que la gente crea en ella, pero debe reconocer que esto no puede alcanzarse por la fuerza. Lo más que puede alcanzar la fuerza es aquiescencia y conformidad externa. Como Hegel dijera del amo del esclavo, el fanático siempre está decepcionado: lo que quería era reconocimiento, pero todo lo que puede obtener es conformidad.

El escepticismo, la indiferencia y las concepciones de la Iglesia amplia no son la única fuente de lo que estoy llamando práctica de la tolerancia. También se puede asegurar en un equilibrio hobbesiano, bajo el cual la aceptación de un grupo por el otro es lo mejor que cualquiera de ellos puede obtener. Desde luego, esto no es en sí mismo una solución de principios, opuesta a la perspectiva escéptica que, a su propio modo, sí lo es. La

solución hobbesiana es además muy inestable. Una secta que pudiera casi poner en vigor la conformidad, podría limitarse a pensar cómo serían las cosas si el otro partido tuviera la preeminencia. Pero para que esto sea un pensamiento hobbesiano, opuesto a un argumento de reversión de papeles que, por ejemplo, se refiera a los derechos, debe existir alguna inestabilidad. Los partidos que son conscientes de tal situación, probablemente buscarán dar golpes preventivos, y tanto más si reflexionan que incluso si sólo pueden esperar aquiescencia y conformidad externa en la generación presente, es concebible que puedan esperar más en generaciones posteriores. De hecho, en el mundo moderno, la imposición por la fuerza de credos e ideologías políticas no ha sido muy efectiva a lo largo del tiempo: una lección que ya era obvia en 1984 era la falsedad, en ese aspecto, del *1984* de Orwell. Sin embargo, la imposición ideológica ciertamente ha funcionado en el pasado, y la especificación de la proposición previa, “en el mundo moderno”, es extremadamente importante. Más adelante regresaré a este punto.

Entre tanto, la tolerancia como *valor* apenas ha aparecido en nuestro argumento. Podemos tener prácticas de tolerancia reforzadas por el escepticismo, la indiferencia o, nuevamente, por un balance del poder sobreentendido. La tolerancia como valor parece exigir más que eso, algo que pueda ser expresado en una cierta filosofía política, una cierta concepción del Estado.

Hasta cierto punto, es posible que las personas pertenecieran a comunidades vinculadas por convicciones compartidas (por ejemplo, religiosas), y que la tolerancia fuera sustentada por una distinción entre aquellas comunidades y el Estado. El Estado no sería identificado con ningún conjunto de tales creencias, y no pondría en vigor ninguna de ellas; igualmente, no permitiría que ninguno de los grupos impusiera sus creencias a los otros, aunque cada uno de ellos podría, desde luego, propugnar lo que cree. En Estados Unidos, por ejemplo, hay un amplio consenso a favor de que la Constitución no permita ninguna ley que ponga en vigor o incluso aliente ninguna religión particular. Hay muchos grupos religiosos, y sin duda muchos de ellos tienen convicciones profundas, pero ninguno de ellos quiere que el Estado suprima a los otros, o que permita a ninguno de ellos suprimir a los demás.

Muchas personas esperan que esto sirva como un modelo general del modo en que una sociedad moderna puede resolver las tensiones de la tolerancia. Por un lado, hay profundas convicciones que difieren sobre asuntos morales o religiosos, sostenidos por varios grupos dentro de la sociedad. Por otro lado, hay un Estado supuestamente imparcial, que afirma los derechos de todo ciudadano a una consideración igual, incluyendo un derecho igual a formar y expresar sus propias convicciones. Éste es el modelo del *pluralismo liberal*. Puede ser visto como representante de la tolerancia. Expresa la peculiar combinación de convicción y aceptación que tiene la tolerancia, encontrando un espacio para las varias convicciones de los diferentes grupos o comunidades dentro del Estado, mientras que la aceptación de la diversidad está ubicada en la estructura del Estado mismo.

Esto no quiere decir que no haya necesidad de creencias compartidas. Claramente debe haber una creencia compartida en el sistema mismo. El modelo de una sociedad que

está unida por una estructura de derechos y una aspiración hacia el respeto en la igualdad, más que por un cuerpo compartido de convicciones sustantivas más específicas, exige un ideal de ciudadanía adecuado para cargar tal peso. La versión más impresionante de ese ideal es quizá la ofrecida por la tradición de la filosofía liberal que fluye desde Kant, la cual identifica la dignidad del ser humano con la autonomía. Una persona libre es la que construye su propia vida y determina sus propias convicciones, y se debe usar el poder para hacer esto posible, no para frustrarlo imponiendo un sistema dado de convicciones.

Éste no es un ideal puramente negativo o escéptico. Si lo fuera, no podría ni siquiera esperar tener el poder de unir en una sociedad a personas con convicciones fuertemente divergentes; ni podría proveer el poder motriz que todas las sociedades tolerantes necesitan para luchar, cuando fallan otros medios, en contra de la intolerancia. Este ideal está asociado con muchos pensadores liberales contemporáneos, como John Rawls, Thomas Nagel y Ronald Dworkin.[93]

Dentro de la filosofía del pluralismo liberal, la tolerancia emerge como una doctrina de principios, y es representada como un valor; más exactamente quizá, como estrechamente relacionada con un valor más fundamental, el de la autonomía. Debido a que se da por sentado que este valor es entendido y compartido, esta explicación del papel de la tolerancia en el pluralismo liberal parece ampliamente justificada. Esto debería proveer un argumento que podrían aceptar aquellos que encuentran perspectivas *prima facie* intolerables en la sociedad, y para suprimir las cuales el liberalismo se rehusa a desplegar el poder del Estado. Thomas Nagel lo ha expresado bien:

El liberalismo es una ideología que justifica la tolerancia religiosa no solamente a los escépticos religiosos, sino al devoto, y la tolerancia sexual no solamente a los libertinos, sino a aquellos que creen que el sexo extramarital es pecaminoso. Distingue entre los valores a los que puede apelar una persona para conducir su propia vida y aquellos a los que puede apelar para justificar el ejercicio del poder político.[94]

Nadie, ni siquiera el mismo Nagel, cree que esto será posible en todos los casos. Debe haber, en cada situación, límites a la medida en que el Estado liberal puede desentenderse de los desacuerdos éticos. Hay algunas cuestiones, como la del aborto, en las que el Estado dejará de ser neutral no importa qué haga. Sus leyes deben trazar distinciones entre diferentes casos de aborto, pero al final no puede escaparse del hecho de que algunas personas creerán con la más profunda convicción que una cierta clase de actos deberían ser permitidos, mientras que otras personas creerán con igual convicción que aquellos actos deberían estar prohibidos. Cuestiones igualmente intratables surgirán con respecto a la educación, donde la autonomía de algunos grupos religiosos fundamentalistas, por ejemplo, para criar a sus hijos en sus propias creencias, será vista por los liberales como algo opuesto a la autonomía de esos niños para escoger las creencias que quieran tener. Ninguna sociedad puede evitar elecciones colectivas y sustanciales en asuntos de ese tipo, y en ese sentido, sobre aquellos asuntos, hay límites a la tolerancia, incluso si las personas continúan respetando las opiniones de los demás.

El hecho de que haya algunos casos como esos, no necesariamente hace naufragar la tolerancia liberal, a menos que sean demasiados. No hay un argumento de principio para

mostrar que si *A* piensa que cierta práctica es errónea y *B* piensa que es correcta, *A* tenga que pensar que el Estado debería suprimir esa práctica o que *B* tenga que pensar que el Estado debería promoverla. Hago estas consideraciones en diferentes niveles; no obstante, hay un famoso argumento a favor de la idea de que el ideal liberal es imposible en principio. Algunos críticos del liberalismo sostienen que el Estado liberal pluralista, como la supuesta puesta en vigor de la tolerancia, no existe realmente. Lo que está sucediendo, dicen, es que el Estado pone sutilmente en vigor un conjunto de principios (*grosso modo*, principios que favorecen la elección individual, la cooperación social y la eficiencia de los negocios), mientras que las convicciones que las personas sostenían profundamente con anticipación, en cuestiones de religión, comportamiento sexual o la importancia de la experiencia cultural, se desvanecen hasta convertirse en gustos privados. De esa manera, el liberalismo sería “meramente otra doctrina sectaria”: la frase que Rawls utilizaba precisamente para explicar en lo que el liberalismo evitaba convertirse.

¿Cuál es la justificación del crítico para decir que el Estado liberal está “sutilmente poniendo en vigor” un conjunto de actitudes en sustitución de otro? Nagel distingue tajantemente entre *poner en vigor* algo como el individualismo, por un lado, y las prácticas de la tolerancia liberal, por el otro, aunque él honesta y acertadamente admite que, por ejemplo, las prácticas educativas del Estado liberal no son “iguales en sus efectos”. Ésta es una distinción importante, y puede hacer alguna diferencia importante en la práctica. No es lo mismo ser sujeto del proselitismo o de la coerción del individualismo militante, que simplemente ver cómo los entornos religiosos tradicionales de uno son erosionados por una sociedad liberal moderna. Los oponentes del liberal deben conceder que hay algo válido en la distinción, pero eso no significa que acepten el uso que hace de ella el liberal, porque no es una distinción que sea neutral en su inspiración. Está asimétricamente sesgada en la dirección liberal, porque se enfoca en una diferencia de procedimiento, mientras que lo que importa a un creyente no liberal es la diferencia de resultados. Dudo que podamos encontrar un argumento de principio que satisfaga las metas más puras y fuertes del valor de la tolerancia liberal, es decir, demostrar que no descansa en el escepticismo o en las contingencias del poder, y que también pueda explicar en principio a las personas racionales, cuyas convicciones más profundas no estén a favor de la autonomía individual ni de los valores relacionados, que se pongan del lado de un Estado que prefiere que decaigan sus valores antes que imponerlos por la fuerza.

Si la práctica de la tolerancia ha de ser defendida como un valor, tendrá que apelar a opiniones sustantivas acerca del bien; en particular del bien de la autonomía individual, y estas opiniones se extenderán al valor y el significado de las características y virtudes personales asociadas con la tolerancia, así como la imposición o no de varias perspectivas sustantivas. Eso no quiere decir que los valores sustantivos de la autonomía individual se hallen mal dirigidos o carezcan de base. El punto es que estos valores, al igual que otros, pueden ser rechazados, y en la medida en que la tolerancia descansa en ellos, también será rechazada. La práctica de la tolerancia no puede basarse en un valor como el de la

autonomía individual, y a la vez esperar escapar de los desacuerdos sustantivos respecto de dicho bien. Esto realmente es una contradicción, porque de una visión sustantiva de los bienes como la autonomía podría desprenderse el valor que se expresa en las prácticas de la tolerancia.

A la luz de esto, podemos ahora entender mejor la imposibilidad o extrema dificultad que parecía tener la virtud personal o actitud de la tolerancia. Se presentó como imposible porque aparentemente requería que alguien pensara que cierta creencia o práctica es absolutamente errónea o mala, y al mismo tiempo que se podía encontrar un bien intrínseco en el hecho de permitir que continuara prosperando. Esto no es necesariamente una contradicción si dicho bien se encuentra no en la continuación de esa creencia, sino en la autonomía del otro creyente. Las personas pueden coherentemente pensar que cierta perspectiva o actitud es profundamente errónea, y que el florecimiento de ésta debería ser tolerado, si ellos también sostienen el valor sustancial de la autonomía o independencia de los otros creyentes. La creencia en la tolerancia como un valor, entonces, no necesariamente implica una contradicción, sino más bien algo familiar, un conflicto de bienes. Sin embargo, esto a su vez da lugar al problema igualmente familiar de que otros no compartan el aprecio a esos bienes; en particular, las personas que el liberal tiene que tolerar probablemente no compartan la visión del liberal del bien de la autonomía, que es la base de la tolerancia, si es que ésta expresa un valor.

Concedido esto, quizá la realidad sea que, como vimos anteriormente, la práctica de la tolerancia no necesariamente descansa en ningún valor. Puede apoyarse por consideraciones hobbesianas acerca de lo que es posible o deseable, o nuevamente por el escepticismo acerca del desacuerdo y su eventual resolución, aunque con el escepticismo, desde luego, se alcanzaría el punto en el que nadie esté lo suficientemente interesado en la cuestión como para que la tolerancia sea, incluso, necesaria. Es importante, también, que estas actitudes no existan en un contexto en el que no haya otros valores. Apelaciones a la miseria y crueldad implicadas en la intolerancia pueden, en circunstancias favorables, tener algún efecto entre quienes no sean partidarios de la tolerancia como valor intrínseco.

Puede ser que la mayor esperanza de la tolerancia se encuentre no tanto en el principio abstracto que presenta el desafío de combinar el máximo del espíritu puro de la tolerancia con la aversión a lo que ha de ser tolerado. Puede más bien basarse en la civilización, o lo que queda de ella, y en su creación más importante, la sociedad comercial internacional. A pesar de enervantes explosiones de fanatismo en muchas direcciones diferentes, todavía es posible pensar que las estructuras de este orden internacional impulsarán el escepticismo acerca de las pretensiones de exclusividad religiosa y de otro tipo, y acerca de los motivos que puedan tenerse para imponer tales pretensiones. Más aún, esto puede estimular tales perspectivas dentro de las mismas religiones. Cuando tal escepticismo se pone al lado de daños generados por la intolerancia, hay una base para la práctica de la tolerancia —una base que es, desde luego, aliada del liberalismo, pero que es menos ambiciosa que el puro principio del pluralismo liberal, el cual descansa en la autonomía—. Se halla más cercana a la tradición

que puede ser trazada desde Montesquieu y Constant, que la fallecida Judith Shklar llamó “el liberalismo del temor”. [95]

Puede ser que las sociedades liberales preserven, en una atmósfera de tolerancia, una variedad de convicciones fuertes sobre asuntos importantes. Sólo el futuro mostrará si esto es así, y también qué tanto importa a la humanidad que gran parte de las convicciones habrán de desvanecerse. Quizá la tolerancia quedará en haber sido un valor provisional, útil en el periodo que va de un pasado en el que nadie había escuchado de él y un futuro en el que nadie lo necesitará. En el presente, es muy obvio que todavía no ha llegado el tiempo en que podamos vivir sin las extrañas prácticas de la tolerancia, pero estas prácticas tienen que ser sustentadas no tanto por un principio muy puro sino por todos los recursos que podamos poner juntos. Al lado de la creencia en la autonomía, dichos recursos son el escepticismo en contra del fanatismo y las pretensiones de sus defensores; la convicción de que la falta de tolerancia acarrea males manifiestos, y, muy ciertamente, el poder.

XIII. El prejuicio humano

Había una vez una perspectiva llamada “humanismo”. En un sentido todavía existe: es un nombre dado estos días a un movimiento de oposición organizada, algunas veces militante, a la creencia religiosa, en particular al cristianismo. Lo que era más o menos el mismo movimiento solía llevar un nombre igualmente heredado del pasado de la filosofía, el “racionalismo”. En Gran Bretaña, las organizaciones ateas, bajo estos dos diferentes nombres, han coexistido, y creo que alguien que escribió incansablemente en los periódicos, pudo haber sido secretario de ambas.

No es el “humanismo”, en ninguno de tales sentidos, de lo que me ocuparé, pero sí haré alguna observación acerca de él, porque es relevante, junto con la idea de humanidad, en nuestra perspectiva ética, que es la cuestión que quiero discutir. El humanismo en el sentido del ateísmo militante encuentra una paradoja inmediata y muy obvia. Su particularidad no es meramente ser ateo —hay todo tipo de modos de serlo—, sino su fe en que la humanidad puede florecer sin religión; más aún, en la idea de que la religión misma es peculiarmente la enemiga de la realización humana. La idea general es que si se pudieran abolir los últimos remanentes de la religión, la humanidad sería liberada y viviría mucho mejor. Pero esta visión se encuentra con el hecho de que, según ella misma, esta cosa mala, corruptora y omnipresente, la religión, es en sí misma una invención *humana*: ciertamente no viene de ninguna otra parte. Así que los humanistas en este sentido ateo debieran preguntarse: si la humanidad ha inventado algo tan desagradable como ellos dicen que es la religión, ¿qué debiera decirles ello acerca de la humanidad? En particular, ¿puede esperarse realmente que la humanidad viva mucho mejor sin ella?

Pero, ése no es el tema. Cuando dije que había una vez una perspectiva llamada “humanismo”, me refería, más bien, a la época del Renacimiento. El nombre se le dio, al principio, a nuevos esquemas de educación, que enfatizaban los clásicos latinos y la tradición retórica, pero vino a aplicarse más ampliamente a una variedad de movimientos filosóficos. Había un interés incrementado e intensificado en la naturaleza humana.[96] Una forma de esto fue la nueva tradición, inaugurada por Petrarca, de escritos acerca de la dignidad y excelencia de los seres humanos (o, como la tradición inevitablemente lo decía, del hombre). Estas ideas ciertamente no eran originales del Renacimiento. Muchos de sus argumentos ya eran familiares, por ejemplo el argumento cristiano de que la superioridad del hombre había sido mostrada por la elección de un ser humano como vehículo de la encarnación, o la más antigua idea, que se retrotrae al menos a Protágoras,

tal y como éste es presentado por Platón, de que los humanos tienen menos ventajas naturales —menos defensas, por ejemplo— que otros animales, pero que han sido más que compensados de eso por los dones de la razón y la cognición.

Otros, desde luego, adoptan una visión más pesimista de los poderes y las potencialidades humanos. A Montaigne le maravillaba cuán peculiares eran los seres humanos, pero no fue nada entusiasta de las razones por las que lo eran. Pero si las visiones eran positivas y elocuentes, o más escépticas o pesimistas, había una característica que casi todas compartían entre sí, y que compartían también con el cristianismo tradicional, y eso es difícilmente sorprendente, puesto que virtualmente cada persona que en el Renacimiento estaba influida por el humanismo era algún tipo de cristiano. Para empezar, casi todo el mundo creía que los seres humanos se hallaban literalmente en el centro del universo (con las excepciones quizá de Nicolás de Cusa y Giordano Bruno, quienes pensaban que no había centro del universo). Fuera de esa creencia puramente topográfica, había una suposición más básica, de que los seres humanos tenían una medida definida de importancia dentro del cosmos. La mayoría de estas perspectivas suponía que dicha medida era alta, que los humanos eran particularmente importantes en el orden de las cosas. Esto es más claro en las versiones más celebradoras del humanismo, de acuerdo con las cuales los seres humanos son los seres más perfectos de la creación, pero se halla presente también en perspectivas que le asignan a los seres humanos una condición maltrecha e imperfecta —la visión de Lutero, por ejemplo, en la que el hombre está lamentablemente caído y no puede hacer nada para remediarlo con su propio esfuerzo—. En éstas, se mantenía la suposición —desde luego, es algo más que una suposición; es más bien una creencia central en la estructura— de que el ser humano es de importancia absoluta. El cosmos puede no estar mirando a los seres humanos, en su estado caído, con mucha admiración, pero ciertamente está mirándolos. La condición humana es una preocupación central para Dios, tan central, de hecho, que condujo a la encarnación, la cual juega su papel tradicional también en la Reforma, señalando el papel especial del hombre dentro de este esquema. Si el destino del hombre es una preocupación muy especial para Dios, no hay nada más absoluto que eso: es una preocupación central, punto.

Las visiones abiertamente antropocéntricas del cosmos desde luego son menos comunes hoy día que entonces. Hace mucho tiempo que la gente ha estado impresionada por el reacomodo meramente topográfico del universo —dejo a un lado las preocupaciones sobre la tierra misma; regresaré sobre este punto—; no estamos en el centro de nada interesante: nuestra ubicación en la galaxia, nada más para empezar, parece casi extravagantemente no comprometedor. Más aún, muchas personas suponen que hay criaturas vivientes en otros planetas de esta galaxia, en otras galaxias, quizá en otros universos. Parece pretencioso o simplemente tonto suponer que esta empresa tiene que ver con nosotros. Incluso cristianos, o muchos de ellos, se impresionan menos con la idea de que Dios debe estar *más* preocupado por los seres humanos de lo que lo está con cualquier otra criatura (me temo que no sé cuál es el estado de pensamiento actual acerca de la encarnación). La idea de la importancia absoluta de los seres humanos parece

muerta, o a punto de estarlo.

Sin embargo, aquí necesitamos ir con algo de cuidado. La suposición que estoy considerando, como la formulé, es que los seres humanos tienen una medida definida de importancia en el cosmos. Dentro de esa suposición, naturalmente, se ha creído que esa medida es alta, y he sugerido que eso en sí mismo puede, a su vez, adoptar dos diferentes formas: la petrarquista o celebrante, en la cual el hombre es espléndidamente importante, y la que podemos llamar forma luterana, en la que lo que es de importancia final es el hecho de que el hombre yace lamentablemente caído. Pero hay otra aplicación menos obvia de la misma suposición: que los seres humanos tienen una medida definida de importancia en el esquema, pero que es muy baja. Conforme a esta visión, los seres humanos tienen una importancia para el cosmos, pero es evanescentemente pequeña. Esto puede no ser una verdad muy emocionante acerca del cosmos, contrastada con las otras perspectivas que mencioné, pero todavía pretende ser una verdad acerca del cosmos; más aún, pretende ser una verdad emocionante, o al menos importante, acerca de los seres humanos. Pienso que esto puede haber sido en lo que Bertrand Russell estaba pensando cuando, por ejemplo, en un ensayo significativamente llamado *El culto de un hombre libre*, escribió acerca de la transitoriedad de los seres humanos, la pequeñez de la tierra, la expansión vasta y sin piedad del universo, etc., en un estilo de retórica de autoconmiseración y al mismo tiempo autoglorificante que hizo que Frank Ramsey dijera que él mismo se hallaba mucho menos impresionado que algunos de sus amigos por el tamaño del universo, quizá porque pesaba 240 libras.

Esta perspectiva puede hacer que las personas sientan que las actividades humanas son absurdas, porque les damos una importancia que realmente no tienen. Si alguien siente de este modo, no tiene mucho caso decirle que en sus creencias hay una confusión: sus sentimientos provienen probablemente de algún sitio que ese comentario no alcanzará. Pero la confusión existe: entre pensar que nuestras actividades no pasan alguna prueba de importancia cósmica, y (contrastado con eso) reconocer que no hay tal prueba. Si no hay tal cosa llamada punto de vista cósmico, si la idea de importancia absoluta en el esquema de es una ilusión, una reliquia de un mundo que no ha sido todavía desencantado enteramente, entonces no hay otro punto de vista mas que el nuestro, en el cual nuestras actividades pueden tener o carecer de importancia. Quizá, de cierto modo, eso es lo que Russell quería decir, pero su jornada a través del carácter de la soledad y la insignificancia, experimentada desde un punto de vista no existente, sólo podía generar el tipo de confusión llamado sentimentalismo. Nietzsche, en contraste, acertó cuando dijo que una vez hubo una estrella en una esquina del universo, y un planeta en torno a esa estrella, y sobre ella algunas criaturas listas que inventaron el conocimiento; entonces murieron, la estrella se fue, y fue como si nada hubiera sucedido. [97]

Por supuesto, hay en principio una tercera posibilidad, entre un punto de vista cósmico y nuestro punto de vista, una posibilidad familiar en la ciencia ficción: que un día podríamos encontrar otras criaturas que tendrían un punto de vista sobre nuestras actividades —un punto de vista que, es muy vital añadir, pudiéramos respetar—. Quizá

la ciencia ficción no ha hecho un uso muy interesante de esta fantasía, pero puede haber algo que aprender de ella, y retornaré a la misma al final de estas observaciones.

Supongamos que aceptemos que no puede plantearse si los seres humanos y sus actividades son importantes o no desde un punto de vista cósmico. Eso no significa que no hay ningún punto de vista desde el cual sean importantes. Ciertamente lo hay, a saber, el nuestro: y no es sorprendente, puesto que el “nosotros” en cuestión, el “nosotros” que somos los que planteamos esta pregunta y las discutimos con otros que esperamos escucharán y replicarán; somos, desde luego, seres humanos. No es nada sorprendente que este “nosotros” frecuentemente se muestre dentro del *contenido* de nuestros valores. Si una criatura es un ser humano o no, es algo que hace una gran diferencia, la mayoría de las veces, en los modos en que tratamos a esa criatura o al menos pensamos que deberíamos tratarla. Dejemos de lado por el momento las distinciones que son fuertemente replicadas por algunas personas, como el asunto de qué es lo que deberíamos comer. De manera menos contenciosa, hablamos, por ejemplo, de “derechos humanos”, y eso significa los derechos que tienen ciertas criaturas porque son seres humanos, y nada más por serlo. Hablamos de “valores humanos”. Desde luego, en Princeton hay un centro de valores humanos. Claro que eso podría referirse solamente a que los valores en cuestión son poseídos por seres humanos, pero en ese sentido puramente posesivo, el término difícilmente estaría añadiendo mucho, puesto que al menos sobre este planeta no hay ninguna otra criatura que tenga valores cualesquiera o, ciertamente, un centro para estudiarlos y promoverlos. Los valores humanos no son sólo valores que tenemos, sino valores que expresan nuestra humanidad, y estudiarlos es estudiar lo que valoramos en tanto que somos lo que somos, es decir, seres humanos.

Ahora hay algunas personas que suponen que si de alguna manera privilegiamos a los seres humanos en nuestro pensamiento ético, si pensamos que lo que sucede a los seres humanos es más importante que lo que sucede a otras criaturas; si pensamos que los seres humanos como tales reclaman siempre nuestra atención y cuidado en situaciones en que otros animales tienen menos o ningún reclamo sobre nosotros, estamos implícitamente regresando a la creencia en la absoluta importancia de los seres humanos. Dichas personas suponen que estamos en efecto diciendo, cuando hacemos estas distinciones entre seres humanos y otras criaturas, que los primeros son, sin discusión, más importantes que las segundas. Esa objeción es simplemente un error. No tenemos que estar diciendo nada de eso, en absoluto. Estas acciones y actitudes no expresan necesariamente más que el hecho de que los seres humanos son más importantes para *nosotros*, lo cual no es sorprendente.

Esa errónea objeción adopta la forma del siguiente reclamo: al privilegiar a los seres humanos en nuestro pensamiento ético, estamos diciendo *más* de lo que deberíamos: estamos reclamando su absoluta importancia. Hay una objeción diferente, la de que estamos diciendo *menos* de lo que necesitamos decir: que necesitamos una razón para estas preferencias. Sin una razón, prosigue la objeción, dicha preferencia será sólo un prejuicio. Si hasta aquí hemos dado alguna razón, no es en favor de estas preferencias; es simplemente la que expresamos al decir “es un ser humano” o “son humanos” o “es

uno de nosotros”, y eso, dicen los objetores, no es una razón. Nos recordarán los prejuicios paradigmáticos, el racismo y el sexismo. ”Porque es blanco”, “porque es varón” son inadmisibles como razones, aunque pueden ser relevantes en circunstancias muy especiales (el género cobra importancia en el caso de emplear a una persona para trabajar en el baño, por ejemplo, aunque incluso podría pensarse en algunos círculos que eso manifestaría también un prejuicio ulterior). Si las supuestas razones de raza o género son ofrecidas sin una explicación, la respuesta que provocan es “¿qué tiene que ver con eso?”. Esas supuestas razones son similares a la forma “él es uno de nosotros”, para un más estrecho “nosotros”. El privilegio humano es en sí mismo otro prejuicio, dicen estos objetores, y tienen un nombre convenientemente desagradable para esto: el “especismo”.

¿Qué tan buena es esta objeción? ¿Con qué exactitud funciona? Tomará un poco responder a esas preguntas, porque requieren que tratemos de ser un poco más claros acerca de las relaciones entre nuestra humanidad, por un lado, y nuestro dar y entender razones, por el otro, y el camino que lleva en esa dirección requiere varias paradas. Pienso que un buen lugar para empezar, es éste: no muchos racistas o sexistas han supuesto realmente que una escueta referencia a la raza o al género —simplemente decir “es negro” o “es una mujer”— constituyera una razón. Se hallarían, por así decirlo, en una etapa previa o bien posterior a eso. En una etapa anterior, simplemente tenían una práctica a duras penas articulada de discriminación: procedían simplemente así y no necesitaban decir nada a sus compañeros que pensaban como ellos y consideraban que dicho comportamiento estaba justificado. Pero llegó el día en que tuvieron que decir algo justificándose: ante aquellos a quienes se discriminaba, si no les podían decir simplemente que se callaran, a los foráneos o a los radicales, o a ellos mismos en aquellos momentos cuando se preguntaban cuán defendible podía ser dicha práctica, y *entonces* tuvieron que decir *más*. Las meras referencias a la raza o al género no satisfarían lo que se necesitaba entonces: igualmente, las referencias a fuentes sobrenaturales que decían la misma cosa no se pudieron sostener por mucho tiempo. Las cosas que al menos parecían relevantes para el asunto —oportunidades de empleo, la franquicia o cualquier cosa que haya sido— eran entonces argumentadas, como la supuesta debilidad intelectual y moral de los negros o de las mujeres. Éstas eran razones en el sentido de que tenían al menos, hasta cierto punto, la forma correcta para serlo, aunque eran, desde luego, muy malas razones, tanto porque eran falsas como porque eran los productos de una falsa conciencia, que operaba para controlar el sistema, y no se necesitaba ninguna teoría social o psicológica muy elaborada para demostrarlo.[98]

En el caso del supuesto prejuicio humano, no parece ser totalmente así. Por una parte, no es simplemente un asunto de discriminación inarticulada o inexpresada: no es un secreto que estemos en favor de los derechos humanos, por ejemplo. Por otra parte, “es un ser humano” parece ser una razón, pero sin la ayuda de ninguna ulterior investigación de razones supuestamente más relevantes, de las que en los otros prejuicios resultaron ser racionalizaciones. Todos estamos conscientes de algunas notables diferencias entre los seres humanos y otras criaturas sobre la tierra, por las que podemos delimitar un espectro completo de criaturas de las cuales podemos decir con confianza

que son seres humanos. Sin embargo, esas diferencias no parecen figurar en nuestro pensamiento como *justificaciones* para proseguir como lo hacemos. De hecho, en muchos casos es difícil entender cómo lo hemos logrado. Únicos en la tierra, los seres humanos usan lenguajes altamente articulados; han desarrollado en una medida sin paralelo el aprendizaje no genético a través de la cultura, poseen literaturas y tecnologías históricamente acumulativas, etc. Hay, desde luego, muchas disputas acerca de la naturaleza exacta y extensión de estas diferencias entre los nuestros y otras especies. Hay discusiones, por ejemplo, de la medida en que otros primates transmiten habilidades aprendidas, y de si tienen tradiciones locales en este sentido. Pero ese no es el punto: hay, en cualquier caso, un hiato conductual nítido y espectacular entre nosotros y nuestros más cercanos familiares primates. Esto se debe, sin duda, a que otras especies homínidas han desaparecido, probablemente con nuestra ayuda. Pero ¿por qué deberían las consideraciones acerca de estas diferencias, verdaderas como ellas son, desempeñar papel alguno en un argumento acerca del vegetarianismo, por ejemplo? ¿Qué tiene todo eso que ver con que los seres humanos se coman a otros animales, pero no a los seres humanos? Es difícil ver argumento alguno en esa dirección que no sea simplemente que el florecimiento de la cultura, la inteligencia y la tecnología es mejor, presumiblemente, que todas esas otras cosas sorprendentes que hacen otras especies del mundo. O considérese, no el caso de comer carne, sino de los insecticidas: si tenemos razón para usarlos, ¿debemos sostener que es simplemente mejor que debiéramos florecer a expensas de los insectos? Si existe algún desarrollo evolutivo espectacular y sorprendente, es la proliferación y diversificación de los insectos. Algunos de ellos son dañinos para los seres humanos, su comida o sus artefactos, pero son verdaderamente maravillosos.[99] Lo que estos últimos puntos muestran es que incluso si pudiéramos apropiarnos de la idea de que es *meramente mejor* el florecimiento de un tipo de animal que el de otro, no es en lo más mínimo claro por qué debiéramos de ser nosotros. Pero el punto básico, por supuesto, es que nosotros no podemos apropiarnos de esa idea en absoluto. Ésta es simplemente otra recurrencia de la noción que vimos hace poco, la importancia absoluta, la última reliquia del mundo todavía encantado. Por supuesto, podemos decir, con razón, que estamos a favor del desarrollo cultural, etc., y pensamos que es muy importante, pero eso mismo es otra expresión del prejuicio humano contra el cual supuestamente estamos luchando.

Así que hay algo oscuro en las relaciones entre la consideración moral “es un ser humano” y las características que distinguen a nuestra especie de otras criaturas. Si hay un prejuicio humano, es estructuralmente diferente de los otros prejuicios de los que ya he hablado, el racismo y el sexismo. Eso no necesariamente muestra que no es un prejuicio. Algunos críticos dirán, por el contrario, que muestra cuán profundo es el prejuicio, hasta el punto de que no podemos ni siquiera articular razones que, pudiera suponerse, lo apoyen. Y si, como dije, parecemos muy dispuestos a profesarlo, el crítico dirá que esto muestra cuán desvergonzadamente prejuiciados somos, o que si podemos profesarlo es porque, muy significativamente, no hay nadie frente a quien tuviéramos que justificarlo, excepto unos cuantos reformadores que son seres humanos como nosotros.

Eso es ciertamente importante. Otros animales son buenos en muchas cosas, pero no para pedir o entender justificaciones. Grupos humanos oprimidos llegan a la mayoría de edad en la búsqueda de emancipación cuando hablan por sí mismos, y ya no a través de miembros reformadores del grupo opresivo, pero los otros animales nunca llegarán a la mayoría de edad: los seres humanos siempre actuarán como sus depositarios. Eso está conectado con el punto, al cual regresaré, de que la única cuestión moral para nosotros en relación con ellos es cómo deberíamos *tratarlos*.

Alguien que habla vigorosamente en contra del especismo y el prejuicio humano es, por supuesto, el Profesor Peter Singer. (Incidentalmente, él tiene su puesto en el Centro para los Valores Humanos de Princeton, que ya he mencionado, y me pregunto cómo interpreta él ese nombre. En el sentido puramente posesivo y cojo de la expresión, es presumiblemente correcta, pero en el sentido más rico, que es seguramente el que le da dicho profesor, debería haber pensado que, de acuerdo con sus valores, que sonaría más bien como un Centro para los valores arios). No importa cuál pueda ser exactamente la estructura del prejuicio humano (si es que es un prejuicio), el trabajo de Singer ha mostrado claramente algunas consecuencias de rechazar que lo sea; consecuencias que él ha estado dispuesto a defender de una forma muy enérgica.

Una idea central implicada en el supuesto prejuicio humano es que hay ciertos asuntos en los que las criaturas son tratadas de un modo más que de otro simplemente porque pertenecen a cierta categoría, la especie humana. A este nivel inicial básico, no necesitamos saber más acerca de ellas. Si se nos dice que hay seres humanos atrapados en un edificio en llamas, solamente por la fuerza de ese hecho movilizamos tantos recursos como podemos para rescatarlos. Cuando se rechaza el prejuicio humano, dos cosas se deducen, como lo ha dejado en claro Singer. Una es que se introduce un conjunto más sustancial de propiedades, supuestamente mejor adaptadas para dar una razón. La segunda es que los criterios basados en esas propiedades, los criterios que determinan lo que se puede propiamente hacer a una criatura, se aplican en cada caso: siempre se trata de que un individuo se ajuste a determinados criterios.

Considérese la cuestión, no de proteger, sino de matar. Singer piensa que nuestras razones para estar menos dispuestos a matar seres humanos que otros animales —la “gran seriedad” de matarlos, como él lo dice— se basan en “nuestros poderes mentales superiores —nuestra autoconciencia, nuestra racionalidad, nuestro sentido moral, nuestra autonomía o alguna combinación de éstos. Son las cosas, estamos inclinados a decir, que nos hacen ser ‘únicamente humanos’. Para ser más precisos, son las cosas que nos hacen personas”.[100]

En otra parte cita con autorización la definición de personas de Michael Tooley, como “aquellos seres que son capaces de verse como egos continuados —esto es, seres autoconscientes que existen en el tiempo—”.[101] Son estas características a las que debiéramos referirnos cuando decidimos qué hacer, y en principio deberíamos tomar cada caso por separado. “Si estamos considerando si es erróneo destruir algo, seguramente deberíamos mirar sus características reales, no sólo la especie a la que pertenece”, y “real” aquí es entendido de tal manera que no deja lugar a la potencialidad.

No puede decirse que un embrión deba obtener una protección especial por ser potencialmente una persona, pues aún no lo es, precisamente como (en la terminología de Tooley) alguien que sufre demencia senil aguda es una ex persona.[102]

Como ya he dicho, Singer saca a colación muy claramente estas dos consecuencias de su visión y se apoya en ellas para alcanzar varias conclusiones controversiales. Me estoy ocupando de la visión misma, el rechazo del prejuicio humano, más que con los detalles particulares de la propia posición de Singer, pero hay un punto que debería mencionar para dejar en claro lo que se halla en cuestión. Lo que rechaza Singer no es exactamente la forma del prejuicio humano a la cual yo y muchas otras personas estamos apegados. Singer considera el siguiente silogismo familiar:

Todo ser humano tiene derecho a la vida.

Un embrión humano es un ser humano.

Por lo tanto, el embrión humano tiene derecho a la vida.[103]

Sería mejor que todos estuviéramos de acuerdo en que la conclusión se deduce de las premisas. Quienes se oponen al aborto y a la investigación destructiva de embriones suelen pensar que ambas premisas son verdaderas. Los que, bajo ciertas circunstancias, los apoyan, deben rechazar el argumento, y por lo común niegan la segunda premisa. Singer niega la primera. Más estrictamente, él piensa que la primera es correcta tan sólo si “ser humano” es tomado como “persona”, pero en *ese* sentido la segunda premisa es falsa, porque el embrión todavía no es una persona. Hay un sentido en el que la segunda premisa es verdadera (el embrión pertenece a la especie), pero en ese sentido de “ser humano” no es verdad que todo ser humano tiene derecho a la vida. Menciono esto porque distingue a Singer de quienes (como los activistas más moderados a favor de la despenalización del aborto) aceptan que el embrión, obviamente, es humano en el sentido de que es un *embrión humano*, pero no aceptan que sea todavía un ser humano, más de lo que un embrión bovino es una vaca. Jonathan Glover una vez causó una furia casi mortal en un distinguido abogado de “provida” por lo que me pareció una observación enteramente razonable: si a este caballero se le hubiera prometido una cena de pollo, y se le hubiera servido un omelet hecho de huevos fertilizados, ¿se hubiera quejado? El punto es importante. La visión estándar, la visión que Singer ataca, es que “ser humano” es una noción moralmente importante aun cuando sólo significa un animal perteneciente a una especie particular, la nuestra; pero aquellos que mantienen esta visión no están comprometidos a pensar que un ovario fertilizado sea ya tal animal, más de lo que lo es en el caso de otras especies.

Pienso que ésta y algunas otras peculiaridades de la posición de Singer provienen en parte de su preocupación por un tipo de controversia: está tratando de combatir las políticas conservadoras que se basan en una noción particular, la santidad de la vida humana. Esto ayuda a explicar por qué su posición sobre el aborto y el infanticidio es la misma que la posición “provida”, pero de cabeza: tanto los partidarios de ésta como él argumentan que “si hay aborto, entonces hay infanticidio”, pero ellos lo toman como una

objeción, y él no. En contra de esto, es muy importante decir que uno puede creer que la noción de ser humano sea central en nuestro pensamiento moral sin comprometerse con el conjunto total de reglas tradicionales que se hallen bajo la etiqueta “la santidad de la vida humana”. [104]

La cuestión más básica, sin embargo, es la planteada por la estructura general de la posición de Singer, y es el mismo tipo de pregunta que ya hemos encontrado. ¿Por qué las sofisticadas propiedades que hacen que una criatura sea persona son “moralmente relevantes” para explicar por qué se puede destruir cierto tipo de animal, mientras que la propiedad de ser un ser humano no lo es? Una respuesta podría ser que favorecemos y estimamos esas propiedades, alentamos su desarrollo, y nos llena de odio y resentimiento que sean frustradas, y esto es difícilmente sorprendente, puesto que nuestra vida entera, y no solamente nuestros valores sino el hecho mismo de que tengamos valores, implican que tengamos nosotros mismos estas propiedades. Buena respuesta, pero no resuelve esta duda, puesto que nosotros también, en una relación compleja con todo eso, usamos la idea de ser humano en nuestro pensamiento moral, y trazamos una línea alrededor de la especie humana que distingue lo que estamos éticamente dispuestos a hacer. Una respuesta diferente sería que es *simplemente mejor* que el mundo perciba las sofisticadas propiedades que nos hacen ser personas, y no sólo que los seres humanos como tales florezcan. Pero esto es, una vez más, nuestro ahora familiar amigo, la importancia absoluta, ese sobreviviente del mundo encantado que trae consigo el pensamiento igualmente familiar y alentador de que las propiedades que poseemos —bueno, la mayoría de nosotros, sin contar a los infantes, los enfermos de Alzheimer, y algunos otros— están siendo vitoreadas por el universo.

Debo decir ya que ésta no es la respuesta de Singer a la pregunta que hemos planteado. Él es un utilitarista, y piensa (hablando muy a grandes rasgos) que la única cosa que importa en última instancia es cuánto sufrimiento hay. En la medida en que deberíamos dar atención especial a las personas, esto sería explicado por el hecho de que éstas muy probablemente estarán expuestas al sufrimiento. No quisiera discutir si esa es una explicación razonable o útil de todas las cosas de las que nos preocupamos en relación con las personas; quiero más bien preguntar otra cosa, que nos lleva de regreso a mi pregunta central sobre nuestras concepciones morales de nosotros mismos como seres humanos viviendo entre otras criaturas. Mi pregunta no es: ¿la visión utilitarista da sentido a nuestras otras preocupaciones en términos de nuestra preocupación por el sufrimiento? Mi pregunta es más bien: ¿Qué tanto da sentido su visión a nuestra preocupación por el sufrimiento mismo?

Muchos utilitaristas, incluido Singer, se inclinan a utilizar el modelo de un observador ideal o imparcial. Un filósofo que propuso una versión de dicho modelo hace cincuenta años describió esta figura memorablemente como “omnisciente, desinteresado, desapasionado, pero, por lo demás, normal”. [105] Existen varias versiones de ese modelo, y en muchas de ellas la figura no es exactamente desapasionada, sino más bien benevolente. Eso puede tener diversos significados, en torno al valor positivo que se atribuye a la satisfacción de preferencias, pero concentrémonos en la más simple

aplicación de la idea: que el observador ideal (OI) esté en contra del sufrimiento y quiera que haya tan poco de él como sea posible. Con su omnisciencia e imparcialidad, él, por así decirlo, asume todo el sufrimiento, no importa cuán exactamente hemos de concebir eso, y lo asume en la misma medida. Se parece mucho, por supuesto, a un representante empequeñecido del Dios cristiano, y esto puede muy bien sugerir que representa aún otra versión del punto de vista cósmico: el sufrimiento o su ausencia es lo que tiene importancia absoluta. Pero supongo que los utilitaristas como Singer esperan que el modelo pueda ser detallado en términos menos fantasiosos.

Ellos despliegan el modelo contra lo que ven como un prejuicio, en particular el prejuicio humano, y la idea detrás de eso es que tenemos el sentimiento, la disposición o la convicción de que el sufrimiento es algo malo, pero expresamos esto de formas irracionalmente restringidas; en formas dictadas por la ley del inverso del cuadrado, en que las distancias pueden ser de todos los tipos: espacial, familiar, nacional, racial o de pertenencia a la especie. Se supone que el modelo del OI corrige esto; si pudiéramos captar todo el sufrimiento como él lo hace, no estaríamos sujetos a estos sesgos parroquiales y nos sentiríamos y actuaríamos de mejor manera. Sin duda la historia de dicha herramienta se halla, de hecho, en una especie de *imitatio Christi* secularizada, y sospecho que algunos de los sentimientos que despierta están conectados con ella, pero los utilitaristas esperan presentarlo como algo independiente de eso; como una herramienta que expresa una corrección racional extensiva de algo que nosotros, desde luego, sentimos.

Así que quiero tomar el modelo en serio: quizá más seriamente, desde cierto punto de vista, que aquellos que lo utilizan. Tengo dos problemas con él. Uno es muy familiar, y tiene que ver con las relaciones entre el modelo y la acción humana. Incluso si pensara que la perspectiva del OI fuera una guía confiable para apuntar a lo que sería un *mejor resultado*, ¿cómo se conecta eso con lo que nosotros —cada uno de nosotros— deberíamos intentar hacer? Con respecto al sufrimiento animal, una forma del problema (una forma que se retrotrae al siglo XIX) es la cuestión de las políticas de protección a la naturaleza. Aunque los seres humanos causen sufrimiento a los animales, directa o indirectamente, buena parte del dolor que éstos reciben es causado por otros animales y eso debe ser visto por el OI. Ciertamente, nos preocupa reducir el daño que nos hacen otros animales, y buscamos en alguna medida reducir el daño que nosotros les causamos a ellos. Eso plantea si no deberíamos procurar también reducir el daño que otros animales se causan entre sí, y generalmente es ése el sufrimiento que tiene lugar en la naturaleza. Los utilitaristas ofrecen algunos argumentos para sugerir que no deberíamos preocuparnos por eso; argumentos que invocan el uso más eficiente de nuestro tiempo y de nuestras energías, pero encuentro difícil evitar el sentimiento de que aquellas respuestas son racionalizaciones pálidas y poco convincentes de una reacción más básica: que hay algo enteramente absurdo en la idea; que representa mal nuestras relaciones con la naturaleza. Algunos ambientalistas, por supuesto, piensan que no deberíamos tratar de mejorar la naturaleza en ese sentido, porque la naturaleza es sagrada y deberíamos interferir en ella lo menos posible, pero ellos, desde luego, no se rigen por el modelo del

OI y su preocupación por el sufrimiento.

Esto conduce a un segundo punto que es fundamental. Hay personas que ven nuestras simpatías selectivas como un filtro tendencioso y con prejuicios del sufrimiento en el mundo; quienes piensan que hay que seguir de cerca, hasta donde podamos, la conciencia del OI, y guiar nuestras acciones por la reflexión sobre lo que asume el OI: me pregunto si ellos alguna vez han considerado cómo sería realmente asumir lo que el OI supuestamente piensa. Sea lo que signifique exactamente ese “asumir”, se supone que implica que los sufrimientos de otras personas y de todas las otras criaturas deberían hallarse vívidamente presentes en nosotros, en algún sentido conectados estrechamente con nuestras motivaciones para la acción, como si fueran nuestros propios sufrimientos o los de las personas por las que nos preocupamos o que se hallan inmediatamente en nuestro entorno. Así es como el modelo supuestamente ha de corregir el sesgo. Pero, ¿cómo tendrían que ser las cosas para que esto fuera así, incluso por unos cuantos segundos? ¿Cómo sería asumir cada sufrimiento por el que esté pasando en un momento dado cualquier criatura? Sería un horror último, una pesadilla insoportable. Y ¿cuál sería la conexión de esa pesadilla con nuestras acciones? En el modelo, el OI supuestamente es un mero observador: no puede hacer nada, pero la idea es que nuestras acciones supuestamente han de tratar de seguir o de ser guiadas por la reflexión sobre lo que él en su omnisciencia e imparcialidad está asumiendo, y si por un momento obtuviéramos algo así como una idea adecuada de lo que eso es, y realmente guiáramos nuestras acciones por ello, entonces seguramente destruiríamos el planeta, si pudiéramos, y si otros planetas con criaturas conscientes contuvieran un dolor similar al que contiene el nuestro, también los destruiríamos.

El modelo tiene cosas completamente al revés. Claro que tenemos razones para aceptar nuestras empatías y extenderlas, no solamente a los grupos más amplios de seres humanos, sino hacia otros animales, en tanto que ellos se hallen en nuestro poder. Esto es ya una disposición humana. La definición del *Oxford English Dictionary* de la palabra “humano” es: “Marcado por la empatía hacia los otros y por la consideración de sus necesidades y aflicciones; sentir o mostrar compasión y ternura hacia los seres humanos y los animales inferiores.”

Podemos actuar inteligentemente a partir de aquellas preocupaciones sólo si las vemos como aspectos de la vida humana. No es un accidente, una limitación o un prejuicio que no podamos preocuparnos igualmente por todo el sufrimiento del mundo: es una condición de nuestra existencia y nuestra salud. Igualmente, no es que las exigencias de la conciencia moral requieran que abandonemos enteramente la vida humana y luego regresemos a regular la distribución de las preocupaciones, incluyendo la propia, mediante criterios derivados de alguna parte. Estamos rodeados por un mundo ante el que podemos reaccionar de varios modos: maravilla, alegría, simpatía, disgusto, horror. Podemos, siendo como somos, reflexionar sobre estas reacciones y modificarlas hasta cierto punto. Podemos pensar cómo se debe administrar esta propiedad humana, y también en su efecto en el entorno, pero es una completa ilusión pensar que esta empresa pueda ser aceptada en algunos sentidos y condenada en otros a partir de otra fuente; una

fuente que no se halle ya relacionada con las peculiaridades de la empresa humana. Es una ironía que esta ilusión, incluso cuando pretende rechazar el así llamado especismo y el prejuicio humano, en realidad comparte ilusiones más antiguas acerca de la existencia de una escala cósmica de importancia en términos de la cual los seres humanos deberían entenderse a sí mismos.

Esta perspectiva, si la miramos a la luz de esas antiguas ilusiones —a saber, la oposición al prejuicio humano— se hallaría más cercana al espíritu de lo que he llamado la versión luterana que al de las versiones celebrantes, en virtud de su afirmación de que los seres humanos están torcidos por su egoísmo. No es como la perspectiva luterana, desde luego, precisamente por su antihumanismo: Lutero pensó que al universo le importaba lo que sucedía a la humanidad, pero de acuerdo con esta visión, todo lo que importa al universo es, a grandes rasgos, cuánto sufrimiento contiene. Pero también hay otra diferencia: Lutero pensó que los seres humanos no se podían redimir sin ayuda, pero los oponentes del prejuicio humano suelen pensar que, con la ayuda de la racionalidad y de estas teorías, pueden hacerlo. En eso hay una semejanza con los así llamados humanistas con los que empecé; los extrañamente optimistas abogados del ateísmo.

He dicho que estar preocupado por el modo en que debieran ser tratados los animales, es parte de una perspectiva humana, o humanitaria, y no he comentado nada que sugiera que no debemos estar preocupados con eso. Pero quiero repetir algo que ya he dicho en algún otro lugar, que, muy significativamente, la pregunta que debemos hacernos es cómo debemos tratarlos.[106] Esto no es verdad en nuestra relación con otros seres humanos, y esto muestra ya que no estamos tratando con un prejuicio como el racismo o el sexismo. Un hombre blanco que piensa que el problema sobre las relaciones entre “nosotros”, como él diría, y otros seres humanos, como las mujeres o las personas de color, es cómo “nosotros” deberíamos “tratarlos” o “tratarlas” a “ellas”, estaría partiendo de un prejuicio, pero en el caso de otros animales, ésta es la única pregunta que podría hacerse.

Así son las cosas aquí, ahora, en este planeta; es una consecuencia del hecho que mencioné antes, que debido a nuestras habilidades para controlar la forma en que actuamos, somos el animal más evolucionado de nuestro entorno. Los seres humanos no tienen trato con ninguna criatura que en términos de discusión, principios, visión del mundo o lo que sea, nos pueda contestar. Pero podría ser de otro modo, y puede ser útil, para cerrar, imaginar algo diferente. Supongamos que, en el modo bien conocido de la ciencia ficción, llegaran criaturas con las cuales nos pudiéramos comunicar hasta cierto punto, que fueran inteligentes y tecnológicamente avanzadas (habrían llegado aquí, después de todo), cuyas relaciones entre sí se hallaran mediadas por reglas convenidas, etcétera. Ahora, hay una pregunta enteramente nueva para el prejuicio humano. Si estas criaturas culturalmente ordenadas llegaran a nuestro planeta, un ser humano que pensara que se trata simplemente de cómo *nosotros* debíamos tratarlas a *ellas*, habría subestimado seriamente tanto el problema ético como, probablemente, el peligro.

El ya fallecido Robert Nozick una vez presentó como argumento a favor del vegetarianismo que si nos arrogáramos el derecho de comer animales menos listos que

nosotros, tendríamos que conceder el derecho a tales visitantes de que nos comieran, si fuesen más listos que nosotros en el grado en que nosotros somos más listos que los animales que comemos.[107] No creo, de hecho, que esto sea un argumento en favor del vegetarianismo, sino más bien una objeción a un argumento para comer carne, y no estoy demasiado seguro incluso de cuán bueno sería para este último (porque el punto del carnívoro podría no ser la distancia entre nuestro nivel de entendimiento y el de los animales, sino el nivel absoluto del entendimiento de éstos). Pero el punto principal es que si se propusieran comernos, sería absurdo debatir sus *derechos*, en absoluto. El filósofo egoísta decimonónico Max Stirner dijo: “El tigre que me ataca se halla en lo correcto, y yo que le respondo también estoy en lo correcto. Defiendo en contra de él no mi *derecho*, sino *a mí mismo*”.[108]

Pero la observación de Stirner tiene que ver con un tigre, y es una cuestión de vida o muerte. Mucho de la ciencia ficción, tal como la pueril película *El día de la Independencia*, define el asunto en aquellos términos desde el principio y es la forma en que facilita las cosas. Es muy fácil, también, si los extraterrestres se hallan aquí sólo para ayudar, en términos que podamos reconocer como ayuda. Los códigos estándar de la ciencia ficción, particularmente en las películas, están diseñados para simplificar tales preguntas. Los hostiles y peligrosos tienden a ser más bien viscosos y desagradables, o rígidos y metálicos (en un ejemplo literario brillante, *La guerra de los mundos* de Wells, son ambas cosas). Los agradables y cooperativos tienen mucho pelo, como el copiloto de *La Guerra de las galaxias*, o son lindos como ET, o hadas etéreas como aquellas cositas en la luz brillante al final de *Los encuentros cercanos del tercer tipo*. Sin embargo, podemos imaginar situaciones en que las cosas serían más difíciles. Las criaturas llegadas podrían ser, desde luego, muy desagradables, por ejemplo, con caras que aparentan ser un hervidero de gusanos, pero si esperáramos lo suficiente para saber qué intenciones tienen, podríamos darnos cuenta de que son muy benevolentes. Simplemente quieren vivir con nosotros —más bien, estrechamente con nosotros. ¿Cómo deberíamos responder a esa propuesta? Podría haber algunos filósofos a la mano para recordarnos la distinción entre los valores morales y los no morales, y para decirnos que su benevolencia y amabilidad son moralmente importantes mientras que el hecho de que sean inolvidablemente desagradables no lo es. Pero supongamos que sus intenciones no agresivas son hacer que el mundo sea, según nuestro criterio, desagradable. ¿Y qué tal si lo desagradable resulta verdaderamente insoportable?

O cambiemos todo e imaginemos otra situación. Los extraterrestres son, de acuerdo con nuestras preferencias, moderadamente bien parecidos, y son, nuevamente, extremadamente benevolentes y razonables; pero han tenido mucha más experiencia exitosa que nosotros gobernando sociedades pacíficas, y han encontrado que necesitan *gobernarlas*, y que demasiada autoaserción de especie o, desde luego, autonomía cultural, muestra ser destabilizadora y destructiva. Así que, sin dolor, nos despojarán ciertamente de nuestros prejuicios, y, hasta donde ello se requiera, de algunas de nuestras peculiaridades culturales y de otro tipo. ¿Cómo debiéramos responder a eso? ¿Querrían los oponentes del especismo que nos uniéramos a ellos —que nos uniéramos a ellos,

desde luego, no sobre la base de que no podemos derrotarlos (lo cual podría ser inteligente aunque no muy heroico), sino por principio?

Esta fantasía nos presenta una situación con aspectos familiares. Es como la de un grupo humano que defiende su identidad cultural, posiblemente étnica, en contra de otro grupo humano que pretende dominarlos o asimilarlos: con la muy grande diferencia, sin embargo, de que, puesto que estamos tratando aquí con otras especies, desde luego, extraterrestres, no se trata de que la variación cultural o étnica esté siendo erosionada por la fusión sexual. (Desde la perspectiva del sexo, debe decirse que, la idea de que el así llamado especismo, el racismo, y una vez más el prejuicio de género son todos similares, parece muy peculiar).

De cualquier manera, la situación imaginaria con los extraterrestres se parece de alguna manera a la situación política familiar. Por mencionar un ejemplo, puede muy bien haber un desacuerdo entre el grupo amenazado, en parte un desacuerdo ético, entre quienes podríamos llamar colaboradores, y otros que se hallen en la resistencia. (Parece que los utilitaristas se unirían a los colaboradores.) En nuestro caso imaginario, los rebeldes se organizarán bajo la bandera “Defiende a la humanidad” o “Levántate por los seres humanos”. Éste es un llamado ético en una disputa ética. Desde luego, eso no convierte a “ser humano” en un concepto ético, o no más que la causa del separatismo vasco; una causa ética, como la de los separatistas vascos, convierte “vasco” en un concepto ético. El concepto ético en cuestión sería algo como: lealtad a, o identidad con, el grupo étnico o cultural propio, y en el caso imaginario, el concepto ético sería: lealtad a, o identidad con, la propia especie. Más aún —y ésa es la principal lección de esta fantasía—, éste es un concepto ético que ya tenemos. Es el concepto ético que se halla en funciones cuando, para perplejidad de los críticos, concedemos una consideración especial a los seres humanos por el hecho de serlo. El hecho de que usemos este concepto implícitamente todo el tiempo explica por qué no hay otro conjunto de criterios que apliquemos a los individuos uno por uno. Es nada más que como están las cosas en la vida real no nos vemos llamados a explicar detalladamente este concepto, porque no hay otra criatura en nuestra vida que pueda usarlo o estar motivada por la misma consideración, pero con una aplicación diferente: es decir, ninguna criatura que pertenezca a otra especie puede articular, reflexionar acerca de, o estar motivada por razones que apelen a la pertenencia a su especie.

Así que la idea de que hay un concepto ético que apela a nuestra pertenencia a una especie es enteramente coherente. Por supuesto, puede haber argumentos éticos acerca de los méritos o el valor de cualquier concepto que apele a algo como la lealtad a un grupo o la identidad con él. Algunas personas, en el espíritu de quienes hubieran sido colaboradores de principio en el caso imaginario, se hallan en contra de tales ideas. En la moralidad política del tiempo presente, la actitud con respecto a esta cuestión es impresionantemente ambigua. Muchas personas, quizá la mayoría de las personas con disposición crítica, parecen oponerse a tales actitudes en grupos dominantes y favorecerlas, hasta cierto punto, en grupos subordinados. (Por qué esto es así, es una buena pregunta, pero no trataré de contestarla aquí.) Otros, nuevamente, respetan estos

conceptos por la motivación y el sentido que le dan a una vida que tiene un carácter rico y particular, en contraste, a extremo, con el ideal utilitario del trabajador social que, con su débil conexión con el OI, va de un lado al otro distribuyendo benevolencia. Al mismo tiempo, sin embargo, aquellos que respetan estas concepciones de lealtad e identidad pueden tener razones para ser escépticos acerca de la retórica coercitiva, las mentiras acerca de las diferencias y la mera violencia que frecuentemente se hallan asociadas con tales ideas y con los movimientos que las expresan. Algunas de estas objeciones se trasladan a los modos en que expresamos la identidad de especie, y es por ello que los oponentes del así llamado especismo y el prejuicio humano muy frecuentemente tienen un punto a su favor acerca de las políticas particulares hacia otros animales, incluso aunque estén equivocados acerca del marco de ideas dentro del cual tales cosas debieran ser condenadas.

Es una buena pregunta si el prejuicio humano, si uno lo quiere llamar así, es ineludible en última instancia para nosotros. Regresemos una vez más a la fantasía de la llegada de los extraterrestres dominantes pero benevolentes, y el consecuente debate entre los colaboradores y los rebeldes. En ese debate, incluso los colaboradores tienen que usar un discurso humanamente inteligible, argumentos que sus prójimos humanos puedan reconocer. Pero ¿implica eso que sus argumentos tendrían que ser exclusivos de los seres humanos? Si fuera así, su situación sería, desde luego, paradójica. Sería como si en las discusiones políticas similares acerca de, digamos, la identidad cultural de los vascos, incluso los asimilacionistas tuvieran que utilizar solamente argumentos exclusivos de la cultura vasca. Así que supongamos que no implica esto necesariamente. La alternativa en el caso imaginario sería que los colaboradores usaran argumentos que compartiera no solamente con sus prójimos humanos, sino también con los extraterrestres. Estos argumentos presumiblemente proveerían la base de su colaboración.

Desde luego, algunos filósofos de la moral piensan que los principios morales correctos son los que pueden ser compartidos con cualesquiera agentes racionales y reflexivos, no importa lo distintos que sean en otros sentidos. Pero incluso si esto fuera así, es importante decir que no sería necesario favorecer a los colaboradores. Esto es porque aquellos principios no nos dirían necesariamente ni a nosotros ni a los extraterrestres cómo compartir una vida juntos.[109] Quizá tendríamos que ser demasiado diferentes en otros aspectos para que eso fuera posible —recordemos a los extraterrestres desagradables— y lo mejor que pudiéramos hacer fuera establecer un pacto de no agresión y coexistencia a distancia. Eso dejaría nuestras peculiaridades —nuestros prejuicios, si es que eso es lo que son— donde estaban, pero supongamos que hemos de vivir juntos. No hay razones para suponer que los principios universales que compartiéramos con los extraterrestres justificaran nuestros prejuicios. Ni siquiera podríamos estar seguros de que justificaran que se nos permitiera tenerlos, por tolerancia; como dije al plantear la fantasía, la larga experiencia y entendimiento benevolente de los extraterrestres podría permitirles ver que tolerar nuestros prejuicios conduce a la inestabilidad y la injusticia, y querrían quitárnoslos, y dentro de esas suposiciones

deberíamos estar de acuerdo con ellos. Los colaboradores deberían entonces estar en lo correcto, porque las concepciones morales que compartirían con los extraterrestres trascenderían las peculiaridades locales.

Pero si esto es así, ¿no se deduce algo más fuerte? Al plantear esta fantasía, dije que el escenario de *El día de la Independencia*, en que los extraterrestres son manifiestamente hostiles y quieren destruirnos, es, para nosotros, un caso éticamente fácil: trataríamos de defendernos. Pero ¿deberíamos hacerlo? Quizá eso sería meramente otra reacción humana irracional, visceral. Los extraterrestres benevolentes, justos y previsores pueden saber mucho acerca de nosotros y nuestra historia, y entender que nuestros prejuicios son irreformables: que las cosas nunca serán mejores en esta parte del universo hasta que seamos eliminados. No estoy diciendo que eso es necesariamente lo que pensarían los extraterrestres informados y benevolentes. Incluso si lo pensara, no estoy diciendo que los moralistas universales, los colaboradores potenciales, tendrían que estar de acuerdo con ellos, pero pueden estarlo, y si estuvieran reticentes, no veo por qué deberían estar seguros de no ser las víctimas de lo que en sus propios términos sería simplemente otro prejuicio que se sirve a sí mismo. Esto, me parece, es un lugar en el que el proyecto de tratar de trascender enteramente los modos en que los seres humanos se entienden a sí mismos y le dan sentido a sus prácticas podría terminar. Y en este punto parece que sólo queda una cuestión: ¿De qué lado está usted?

En muchas conexiones más limitadas, las esperanzas de un automejoramiento pueden encontrarse peligrosamente cercanas al riesgo del odio a uno mismo. Cuando la esperanza es mejorar a la humanidad al punto en que todo aspecto de su dominio del mundo pueda justificarse ante una corte más alta, es probable que el resultado sea o bien el autoengaño, si se piensa que se ha tenido éxito, o bien el odio a uno mismo y el autodesprecio cuando se reconozca que siempre se fracasará. El odio a uno mismo, en este caso, es un odio a la humanidad. Personalmente pienso que hay muchas cosas que detestar de los seres humanos, pero su sentido de identidad ética como especie no es una de ellas.

Tercera Parte

El alcance y los límites de la filosofía

XIV. La filosofía política y la tradición analítica

No hace mucho tiempo la filosofía analítica se dio por vencida, o casi, en la filosofía política. Las introducciones a volúmenes sucesivos de *Politics, Philosophy and Society* expresaban la duda de que el tema pudiera continuar existiendo, o la sorpresa de que aún lo hiciera; al menos un simposio internacional tuvo como título “La philosophie politique, existe-t-elle?” y muchos otros tenían el mismo tema. No es necesario decir que ese tiempo ha pasado ya. No me preocupo por aclararme exactamente qué filosofía es analítica, ni alentando la discusión sobre ese tema tan ingrato, sino que doy por sentado que, por ejemplo, *A Theory of Justice* se halla en la tradición analítica, como lo están *Anarchy, State, and Utopia* de Nozick, y la mayoría de las contribuciones a *Philosophy and Public Affairs*. Claramente, el predicho funeral ha sido indefinidamente pospuesto.

En parte, supongo que el cambio puede ser explicado por la ley —que hasta donde yo sé carece de excepciones, pero que no es transparente para todos— de que la filosofía política viva surge solamente en contextos de urgencia política. La somnolencia de la filosofía política fue, en esa medida, un fenómeno del periodo que prematuramente saludó el fin de la ideología. Pero si esto es así, ya nos dice algo acerca de la incómoda relación de la filosofía política con la tradición analítica: la dependencia habitual, y parece que imprescindible, que la filosofía política tiene con la urgencia de los problemas políticos, que en primer lugar no son filosóficos, forma una sola pieza debido a su insistencia, cuando tiene que hacerlo, con ser tanto normativa como impura. Es normativa al menos en el sentido de que el desacuerdo moral y político de primer orden con el autor, puede motivar de manera importante un desacuerdo con su filosofía, y es impura en el sentido de que los materiales extraídos de fuentes no filosóficas —una relación con la historia o las ciencias sociales, por ejemplo— probablemente tienen un papel más que ilustrativo en el argumento.

La filosofía analítica de aquellos años quiso, y lo intentó seriamente, no ser ni normativa ni impura. La distinción entre hecho y valor (o más bien, en este caso, entre teoría y valor) supuestamente servía para segregar lo filosófico de lo normativo, mientras que la otra distinción entre analítico y sintético servía para segregar lo filosófico de lo histórico o lo referente a las ciencias sociales. Concedido este programa, y concedidas las características hereditarias de la filosofía política, era de esperarse que la filosofía analítica ya no podría seguir siendo realmente útil. Así que la explicación del letargo de la filosofía política en este tiempo no puede simplemente descansar sobre las características de la época; debe además explicar por qué la filosofía política es particularmente reacia a

convertirse en una disciplina pura y no normativa, y por qué la filosofía analítica tenía en ese momento ambiciones de reformarla.

Discutiré la primera pregunta, de manera más bien oblicua, más adelante. No intentaré contestar la segunda, pero bosquejaré lo que una respuesta tendría que explicar. Frecuentemente, se sugiere que las ambiciones reformadoras de la filosofía analítica surgieron exclusivamente por haber aceptado estrictamente las dos distinciones que he mencionado, en particular la de hecho-valor, pero eso debe ser un error, pues incluso concedida tajantemente esa distinción, tiene que agregarse una doctrina del papel propio de la filosofía para determinar que ésta se ocupará de una, pero no de la otra. Este punto está bien ilustrado por la obra de uno de los padres modernos de la distinción hecho-valor en la filosofía analítica, G. E. Moore. Moore, después de haber anunciado en *Principia Ethica* (Cambridge: Cambridge University Press; trad. al esp., México: UNAM, 1959) la existencia de la así llamada “falacia naturalista”, y de declarar que la consecuencia de ésta es que ningún proceso puramente lógico podía conducir a nadie desde proposiciones metafísicas hasta juicios de valor, dedicó sin reservas gran parte de su libro a elaborar sobre lo que, desde su punto de vista, es bueno. Por supuesto, concedida la falacia naturalista, estos comentarios deben ser, en efecto, adicionales a sus aseveraciones metafísicas y lógicas, pero no le importó hacerlas. Difiere de sus sucesores no en sus concepciones de la relación entre metafísica o lógica con las afirmaciones de valor, sino en la idea de que los libros escritos por los filósofos deberòam confinarse a la metafísica y la lógica.

En parte, el origen de esta diferencia se puede encontrar en otra diferencia, acerca de la epistemología de los juicios de valor. Podría haber alguna razón, si bien oscura, para suponer que si el descubrimiento de lo que es bueno descansara, como suponía Moore, en la intuición de propiedades no naturales, entonces un filósofo tendría alguna habilidad apropiada para mantener objetos intelectuales relevantes en una suspensión transparente en su mente —una habilidad no peculiar a la filosofía pero al menos favorecida por ella—, pero una vez que los últimos desarrollos de la teoría hecho-valor condujeron a explicaciones no cognitivas de lo que es sostener juicios de valor, entonces, desde luego, hubo una dificultad para esa pretensión del filósofo, sobre la atención de cualquiera. En particular, la tendencia a considerar juicios de valor (o más bien, decirlos en voz alta) como primordialmente protrépticos, como buscando exhortar u ordenar a sus escuchas, oscureció la relación entre el filósofo que dice tales cosas y su público.

Ha habido un problema real acerca de la relación del moderno filósofo moral o político con su público: qué es lo que éste requiere de la atención de los demás, y habré de regresar a ello, brevemente, al final de este artículo, aunque en un contexto, espero, menos llanamente desalentador que el presentado por el tipo de concepción que acabo de mencionar. Pero en la conexión presente, es importante decir que las limitaciones de la filosofía analítica difícilmente pudieron haber sido impuestas sólo por la distinción hecho-valor; requerían también una concepción especial de las responsabilidades de la filosofía. También hay que decir que las limitaciones de la filosofía, y la asociada sequedad de la filosofía política, no son únicamente alentadas por la creencia, sin lugar a dudas

importante en muchas concepciones del “fin de la ideología”, de que el desacuerdo serio de valor estaba por terminar y que iba a tener lugar el consenso sustancial. Por el contrario, fue precisamente un sentido del contraste entre la pluralidad de valores, y su irresoluble conflicto, como opuesta a la supuesta universalidad de la lógica y la ciencia, lo que ayudó a motivar la distinción entre hecho y valor.

La distinción hecho-valor y cómo verla en tres dimensiones

He mencionado por qué la filosofía política no puede prescindir de su relación tradicional con lo normativo y lo impuro y conservar una alianza pura y productiva con la filosofía analítica, limitada por su ambición de permanecer pura. La respuesta corta es que la operación peculiarmente bidimensional de la distinción hecho-valor, como se empleaba entonces, ofreció todo lo que tenía que ofrecer a la ética o al valor *en general*, y no dejó nada interesante que decir acerca de los asuntos políticamente distintivos. La distinción imponía un contraste entre los elementos del lenguaje que registraban el estado del mundo, y aquellos que expresaban políticas, principios o decisiones para cambiarlo, o en cualquier caso, en otra versión, reacciones afectivas que se relacionaban con deseos de cambiarlo. Muchos de nuestros más interesantes conceptos de valor evidentemente combinaban estas dos funciones, y servían, así, tanto para registrar conjuntos complejos de hechos como para evaluarlos. Pero la evaluación tenía que ser lógicamente separable de los hechos, o cierta forma de describir el mundo podía en sí misma incluir evaluaciones. Un hecho implica valor y, básicamente, cierto tipo de libertad que exige esta concepción de que los valores de un individuo no deberían ser dictados por el mundo.

Si esto fuera cierto, no podría decirse nada de verdadero interés filosófico acerca de estos complejos conceptos de valor. La filosofía habría logrado establecer su tesis general acerca de la separación del elemento valor, y la pregunta de por qué los elementos descriptivos deberían agruparse, y cómo es que las evaluaciones relacionadas con un contexto más amplio de creencias serían dejadas a las ciencias sociales y así, por virtud del requerimiento de pureza, definitivamente fuera de la filosofía. Desde este punto de vista, el complejo concepto de valor no podría sugerir la pregunta de cómo es que ciertos hechos implican ciertos valores. Esa presumiblemente sería una pregunta filosófica, si fuera posible, pero desde esta perspectiva no lo es. Ningún hecho implica valor. Nosotros, u otras sociedades, aplicamos valores a algunos hechos, y las preguntas sugeridas por el complejo concepto de valor pueden ser solamente cómo y cuándo, y a qué hechos nosotros o ellos damos valor, y eso es visto como una pregunta para las ciencias sociales. Pero ahora el tema distintivo de la filosofía política debe implicar desde luego conceptos complejos de valor, pues no es la vieja distinción entre lo correcto y lo incorrecto, sino los temas importados por la legalidad, la justicia, la igualdad, o la libertad, los que son sus preocupaciones. De ahí que la teoría bidimensional hecho-valor no podía encontrar nada de interés para la filosofía política, y no es un accidente que esta última

haya preservado su vieja obstinación en ese punto.

Tampoco es accidental que deba permanecer impura, pues si ha de haber un modo filosófico de hacer las cosas mejor con respecto a la complejidad de estos conceptos de valor y sus relaciones con un trasfondo más amplio, no será uno que deje totalmente atrás los intereses de las ciencias sociales, sino más bien uno que coopere con ellas. Desde luego no quiero tratar de dar aquí sugerencias extensivas acerca de la distinción hecho-valor y lo que debería sucederle, pero vale la pena mencionar una posibilidad que curiosamente no atrajo tanto interés de los teóricos de la distinción hecho-valor como hubiera podido hacerlo: a saber, que no hay, en absoluto, compulsión a usar un concepto dado de valor. La teoría bidimensional hecho-valor implicaba que para cualquier concepto mixto *C* siempre era posible tener un concepto *C'* que tuviera todo (y solamente) el contenido descriptivo de *C* pero sin su poder evaluador, y que sería una pregunta no filosófica si *C'* tiene o no alguna utilidad, pero podemos tomar como una consideración filosófica que *C* perdería tanto su identidad como su sentido sin su poder evaluador, y que su identidad, más aún, puede estar afectada en sus relaciones con un más amplio rango de conceptos. Al enfatizar esta consideración, uno podría esperar reconocer que las evaluaciones pueden estar más íntimamente ligadas con maneras de describir el mundo de lo que suponía la explicación analítica temprana, de manera que, por ejemplo, la selección de ciertos tipos de conducta para evaluación tiene sentido en sí misma solamente dentro de un marco general de creencias, y, no obstante, uno podría al mismo tiempo sostener la verdad de que las creencias morales no pueden ser simplemente un registro de la forma en la que es el mundo, e incluso, con cierta diferencia, preservar la perfecta libertad de los valores que los teóricos de hecho-valor deseaban.

El concepto de *pecado*, por ejemplo, relaciona en sí mismo hecho y valor, y los relaciona de un modo complejo, en el que no son externos entre sí —uno no podría tener todas esas creencias y abandonar esos valores—, pero, como revela la vida moderna, no hay necesidad de que los seres humanos usen el concepto de pecado. Si lo vemos de esa manera, puede recuperarse el sentido tridimensional de un conjunto de creencias acerca del mundo y la sociedad de acuerdo con el cual los valores tienen un significado; se puede examinar la estructura detallada de un conjunto de valores de este tipo sin suponer que es la única posible; es probable aproximarse a los sistemas de valores de otros tiempos y lugares con un conjunto de categorías más realista y flexible, y necesitará, y podrá, usar ventajosamente las ciencias sociales en lugar de considerarlas un almacén de lo que no es filosófico.[110]

Solamente con una diferencia esto cumple con los requisitos del valor máximo de la libertad. La idea de que uno podría, por así decirlo, abandonar el compromiso con un valor particular dentro de un complejo de valores presumiblemente significaría, si significara algo, que uno tendría un tipo particularmente individual de libertad; pero si reconocemos que los valores se hallan más profundamente incorporados en los sistemas de creencias, la libertad en cuestión se vuelve más radical y menos asequible, pues parece que la reconstrucción de una perspectiva se halla menos a la mano. La libertad se

volvería de un hombre más que de los hombres, pero fue solamente la abstracción extrema de la concepción anterior lo que dio la impresión de otra cosa. Es interesante que el exponente más puro del tipo temprano de concepción analítica que he estado discutiendo, R. M. Hare, haya adoptado ahora una posición mucho más normativa, pero es precisamente una posición utilitarista la que preserva el individualismo implícito en la concepción anterior, y, al mismo tiempo, la opinión de que los conceptos de complejos de valores sustantivos son en principio redundantes o carentes de interés.

Comprensión social reflexiva

Este tipo de elemento puede verse por la filosofía analítica como ideológico; no, quizá, en el sentido de que su difusión sirva a un interés, pero sí al menos en que su sentido y presuposiciones se forman de modos no evidentes en la superficie, y quizá no evidentes a los escritores mismos, que admiten la explicación social. No quiero discutir la cuestión de si estas explicaciones son veraces: tienden, de hecho, a moverse entre lo vago y lo anecdótico, pero me gustaría sugerir que en cualquier caso la filosofía analítica, entre todas las filosofías, no ha contado hasta ahora con los elementos necesarios para considerar si estas explicaciones podrían ser ciertas para plantear preguntas reflexivamente acerca de sus propias relaciones con la realidad social. La extrema abstracción a la que ya me he referido, y el carácter conceptual de su tema, no solamente la apartan de un modo obvio de consideraciones de este tipo, sino que en realidad la excluyen. Una reflexión epistemológica puramente conceptual del estatus de las tesis de la filosofía analítica, se halla por supuesto disponible, y ha atraído demasiado la atención, pero la reflexión histórica concreta fue excluida por la proscripción de lo empírico.

En tanto que la posición puramente conceptual promueve esta inmunidad a la reflexión social, uno podría esperar que motivara también una mayor apertura a las impurezas de las ciencias sociales. Claro que podría, en el sentido obvio e inmediato de que mucho de la ciencia social en este momento está obsesionada con tales asuntos, y alguna sociología da la apariencia de haberse derrumbado en epistemología social pura. Pero si uno llegara a creer, en términos generales, que la mera presencia de las ciencias sociales fuera a alentar tal reflexión, sería claramente optimista, puesto que algunas ramas de dichas ciencias, en particular algunos tipos de ciencia política, son las que han dado más lugar a quejas por carecer de tal autocrítica reflexiva. Podría decirse que esta crítica puede evitarse si la idea de comprensión social se uniera a la filosofía ya que ésta es esencialmente reflexiva. Pero suponer que la fusión de la filosofía reflexiva con la conciencia de la realidad social da por resultado conciencia social reflexiva, podría ser la visión optimista de un consejero matrimonial: como cuando Bernard Shaw le contestó a una actriz, “¿y qué tal si resulta con su inteligencia y mi belleza?” Para tomar un ejemplo particular, la teoría de las ciencias sociales de Winch mezcla una apertura a los datos antropológicos con un método filosófico; obviamente representa una asimilación muy cercana del entendimiento social por el conceptual, y ciertamente no está mejor dotado

de conciencia reflexiva que la filosofía analítica o la ciencia social positivista.

De hecho, no hay un modo mecánico de garantizar que los filósofos políticos y morales sean más sensibles a estos asuntos, como deberían serlo, aunque la sensibilidad debería estar dispuesta, si se presentara la ocasión, a mirar la dificultad cara a cara y seguir adelante (reconociendo una tradición de la filosofía de afrontar una dificultad y continuar con algo en lo que uno cree). La lección que extraigo de ello es que la educación de los filósofos políticos debería incluir materiales epistemológicos que les ayudaran a adquirir en cierta medida las diferentes aseveraciones de la sociología del conocimiento. Se ha dicho que los metafísicos y los filósofos del lenguaje no deberían ser verificacionistas, pero sí deberían tener una conciencia verificacionista; así los filósofos políticos deberían tener una disposición a sentirse turbados por la posibilidad de reflexionar sobre la formación y dirección de sus concepciones.

¿Poco a poco o sistemático?

La cuestión de cuán sistemática debiera ser la filosofía, y la pregunta relacionada, si es que no idéntica, en qué medida debiera constar de teorías, ha sido el objeto de muchos desacuerdos dentro de la filosofía analítica; sobre esa cuestión hay, a grandes rasgos, una tradición británica de improvisación por pedazos (con la conspicua excepción de Russell) opuesta a una tradición teutona-americana dirigida a las teorías. La forma en que se concibe hoy esta cuestión corresponde en buena medida a la fortuna que han tenido, poética y económicamente, estos dos agrupamientos, y difícilmente se puede dudar que el enfoque más sistemático, y basado en las teorías, simple y sencillamente ha triunfado en áreas centrales de la filosofía. En la filosofía del lenguaje, se ha afirmado, muy en especial, la tesis de que una distinción o análisis aislados carecen de objeto y de sentido: la concepción profesada de Austin, que consiste en coleccionar distinciones lingüísticas como si fueran tipos de escarabajos, puede ser absurda aplicada a distinciones lingüísticas, y no muy brillante aplicada a escarabajos.

En una teoría del significado como la de Davidson y la de Quine, la noción misma de teoría tiene un papel central: pues se introduce la noción de significado (en tanto que *es* introducida; la tesis vale también con respecto a los sustitutos del significado en estas teorías) esencialmente en términos de qué condiciones están asociadas con una expresión dada por la teoría que óptimamente se ajusta a todas las expresiones observadas. Ahora, a lo que nos referimos aquí es, desde luego, a una teoría empírica, una teoría que un lingüista podría formar acerca de un lenguaje dado que estaba tratando de entender, pero la tesis tiene ulteriores ramificaciones hacia la filosofía misma: pues el filósofo tratará sistemáticamente de dar análisis o elucidaciones de expresiones de nuestro propio lenguaje que serán de un tipo para que encajen, al menos, con tal teoría empírica, y sin mucha restricción la elección entre análisis posibles o elucidaciones de expresiones se vuelve indeterminada y carente de objeto. Este es un ejemplo extremo, pero se puede discutir la misma tesis para abarcar más; es sólo en términos de lo que podría decirse de

muchos casos, o expresiones, o áreas que la elección de qué decir acerca de *este* caso se vuelve determinada. Es la ausencia de estas limitaciones lo que hace gran parte de la filosofía wittgensteiniana tan vacía. La concepción original era que lo que hace que la elección de una elucidación o una observación filosófica tenga las determinaciones apropiadas es una *preocupación*, y lo que hace tan impresionante a mucho del trabajo posterior de Wittgenstein es la presencia de una preocupación reconocible —la suya propia—, pero muchos de sus seguidores parecen abordar preocupaciones postuladas o preocupaciones tipo, que a lo mucho hacen restricciones muy débiles a lo que es, de hecho, una teoría muy confusa.

Las limitaciones sistemáticas a la filosofía del lenguaje valen para otras áreas muy generales de la filosofía: la metafísica, la teoría del conocimiento y la filosofía de la mente, en parte, porque *valen* para la filosofía del lenguaje. Sin algunas de tales restricciones —y no quiero exagerar el grado en el que ya son claras o aceptadas—, es difícil ver por qué una observación filosófica debería ser más relevante que otra, o qué podría admitirse como una explicación. En este sentido, vagamente, como ya lo he señalado, quiere decir seguramente que la filosofía debería ser sistemática. La filosofía moral y la filosofía política son también partes de la filosofía, pero no se deduce de eso que debemos necesariamente tener una filosofía moral y una filosofía política sistemáticas, y si así fuera, sería solamente en un sentido débil, que acepto, que estas ramas de la filosofía deberían responder a demandas sistemáticas provenientes de otra parte: por lo pronto, la gramática de las afirmaciones morales o políticas —la gramática literal, no la wittgensteiniana— debe ser la misma que para otras afirmaciones. Lo que no se deduce es que las filosofías moral y política deberían tener sus propios sistemas, o que el sistema de oferta debería ser la demanda prioritaria a los filósofos en estas áreas.

La razón de lo anterior es que una meta importante, y consecuencia cierta, de sistematizar en estas áreas, es reducir o eliminar el conflicto entre nuestras ideas y nuestros sentimientos, y antes de que nos dispongamos a hacer eso, y mientras que lo hagamos, debemos reflexionar en la importancia del conflicto. El conflicto entre nuestros sentimientos y nuestras creencias morales es, en primer lugar, un fenómeno condicionado histórica, social y, tal vez, psicológicamente; el producto de las sociedades plurales y el rápido cambio cultural, así como, quizá, de necesidades psicológicas más generales que tienden a entrar en conflicto. Podemos, en alguna medida, entender *por qué* tenemos sentimientos en conflicto, pero eso no debería significar que, debido a eso, dejemos de ser leales a ellos. En segundo lugar, no es verdad que cualquier situación en la que haya tal conflicto sea peor que cualquiera otra en la que no lo haya, o incluso —lo que es quizá más plausible— que la reducción de conflicto sea siempre una prioridad. En el caso del conflicto de creencias y de las teorías explicativas, la reducción del conflicto es una meta en la que no se duda: ya sea por la razón pragmática de que la eliminación misma del conflicto es la meta de la empresa explicativa, o por la razón realista de que dos creencias en conflicto no pueden ambas, en un sentido sustancial, ser verdaderas; pero la articulación de nuestros sentimientos morales no necesariamente obedece a estas restricciones, y exigir que sean moldeados por los requerimientos del sistema es alterar

nuestra percepción moral del mundo; no sólo hacerla más racional en algún sentido incontestable.

Esto es lo que hace que el modelo de Rawls sea tan engañoso, porque asimila “intuiciones” morales a las intuiciones de quien se expresa naturalmente tal y como lo dicta el contenido de una teoría lingüística. No es simplemente, para la teoría, que se halle más comprometido en la educación de las respuestas morales, en un “equilibrio reflexivo”, que las razones para descartar finalmente algunas expresiones marginales como indeseables, sino más bien que el papel de la teoría es diferente. Corresponde a la teoría lingüística explicar y predecir expresiones aceptables, y es *la* teoría para hacerlo. Pero en el caso de los sentimientos morales, una teoría rawlsiana, es decir, un conjunto de principios o nociones morales que unifiquen nuestras opiniones morales, y en ese sentido prediga la forma en que podría reaccionarse a un caso posible o imaginario, no se halla sola en el campo, pues podemos entender igualmente una teoría externa, *e.g.* a las explicaciones sociales, que prediga que ningún conjunto de tales principios o nociones lo conseguirá. Así que podemos tener sentimientos sistemáticamente conflictivos, por ejemplo, acerca del valor del carácter y del valor de las acciones e intenciones particulares, o del valor de sentimientos particulares no morales y el valor de sentimientos morales, impersonales, de justicia. Podemos sentir la fuerza de los argumentos tanto utilitaristas como antiutilitaristas. Más aún, quizá podamos ver por qué tenemos estos sentimientos conflictivos, y si ello es así, no podemos estar de acuerdo en que debe haber una teoría unificada que “prediga” la respuesta a varias situaciones, a la manera rawlsiana, o bien que algunas de nuestras respuestas hayan de ser desechadas.

De hecho, hay dos comentarios diferentes que se pueden mencionar aquí. Primero, la mera posibilidad de dar una explicación externa al conflicto de valores significa que la pregunta rawlsiana paralela a la pregunta de la filosofía del lenguaje, “¿Qué conjunto de principios unificado y sistemático habrá de predecir estas reacciones?”, no debe de tener necesariamente una respuesta. La unidad de un lenguaje —incluso, la unidad de *mi* lenguaje— está dada por la teoría explicativa misma; no así la unidad de mis valores o sentimientos, y las explicaciones externas de la manera en que llegué a ellas pueden resultar no unificadas (un hombre entre dos culturas no es un hablante efectivo de un lenguaje criollo).

El segundo punto es que, aun cuando estos conflictos se vuelven claros para una persona, queda en pie la duda de cómo y qué tanto ésta podría reducirlos. Por ejemplo, muchos de nosotros tenemos ahora dos tipos de sentimientos acerca de personas a las que se va a privar de la vida: sentimientos utilitaristas por un lado y, por el otro, sentimientos que tienen una articulación compleja, en los que entran en juego cuestiones como quién ha de ser muerto, o bien, que no hacer ninguna elección es mejor que hacerla, y que, si ha de elegirse, deben hacerse algunas consideraciones estructurales (por ejemplo, que sean los menos posibles, o que es preferible que mueran quienes ya se están muriendo en lugar de los sanos, etc.). Esta conjunción de sentimientos conduce al conflicto, y los utilitaristas en particular argumentan en favor de *a)* una racionalización de nuestra perspectiva, y *b)* que dicha racionalización se haga en la dirección utilitarista,

diagnosticando el segundo conjunto de sentimientos como el residuo de alguna perspectiva no utilitaria anterior; pero hay al menos tres niveles en los que se puede resistir esta presión a la racionalización.

Primero —y esto es algo que los mismos utilitaristas pueden reconocer—, las consecuencias de la realización social efectiva pueden ser más extensas y dañinas que lo que se espera. Segundo, en un nivel psicológico, puede ser más difícil vivir y manejar los valores racionalizados. Y tercero —lejos ahora de preocupaciones utilitaristas—, puede ser que estos sentimientos se hallen metafísicamente más relacionados con nuestra concepción de las personas de lo que parece superficialmente. Eso requiere exploración, y esa exploración requiere paciencia con nuestras aparentes irracionalidades. Y quizá más que paciencia, al menos, paciencia indefinida.

Es una pregunta abierta, y pienso que muy difícil, qué tan incuestionable es el ideal de que, incluso en última instancia, las respuestas morales deban estar armoniosamente integradas, de un modo en que la mayor parte de las nuestras por ahora evidentemente no lo están. (Una pregunta relacionada, igualmente difícil, es a qué grado esta integración existe respecto de la racionalidad en primera persona; tal exigencia, evidentemente muy lejos de ser cumplida, la hacen Rawls, Nigel y muchos otros para decir que tendríamos que planear racionalmente nuestra vida como un todo, como si se nos hubiera dado un rectángulo para iluminar a la perfección.) Pero cualquiera que sea la respuesta de esas cuestiones, es cierto que tal fin ideal no ha de considerarse realizado en la estructura profunda de nuestros sentimientos existentes, y que no es posible aproximarse a él educando dichos sentimientos de acuerdo con el sistema. Primero, tienen que entenderse, lo cual, en el contexto de una filosofía sistemática en general, es probablemente una labor ardua y desordenada.

Puede decirse que la moralidad requiere acción, y que la acción requiere que se tomen decisiones —uno espera, sobre principios racionales—. Es verdad que tenemos que actuar, pero es solamente la concepción de que la acción agota el objeto y contenido del pensamiento moral, lo que puede conducir a alguien a pensar que eso es el fin de este asunto. Lo que sigue, más bien, son hechos familiares, como que a veces actuemos (y actuamos necesariamente) sin una completa convicción, o nuevamente —y esto es menos familiar, que la bien conocida falta de convicción liberal, y quizá más estimulante— que es posible actuar con entera convicción en una ocasión, consciente de que otra vez, en circunstancias casi iguales, la convicción pudo no llegar. Las relaciones entre acción, convicción y racionalidad son un área en que la filosofía moral, la filosofía de la mente, y la metafísica se encuentran de una manera muy significativa.

Sea como fuere, se puede decir razonablemente que la política es un asunto diferente, pues mientras que las personas individuales pueden titubear en diferentes medidas entre sus sentimientos morales conflictivos, hay una exigencia de consistencia racional y principio en las posiciones públicas (muy aparte de lo que está legalmente promulgado). Eso es evidente, pero entre muchas cosas más que deben decirse al respecto, hay dos que debemos mencionar aquí por su efecto en un programa para la filosofía política. Primero, el requerimiento de consistencia de principios que logra en una

serie de decisiones públicas provenientes de una autoridad establecida en un Estado racional (idealizado, por supuesto), o en el programa de un partido político (incluso más idealizado si es posible), obviamente no se extiende a la vida política como un todo, la cual precisamente puede encarnar no solamente conflictos de interés y de principios directamente opuestos, sino de valores en conflicto, y de interpretaciones en conflicto de los mismos valores (considérense aquí los conflictos entre la igualdad de oportunidades y la igualdad de estima, o de la justicia como igualdad con la justicia como derecho). Los pensamientos del filósofo no tienen que estar dirigidos a resolver estas diferencias; puede servir más, de hecho, para agudizarlas, dejando en claro en qué modos ambas están enraizadas en nuestros sentimientos. Esto puede no ser, en el corto plazo al menos, una actividad enteramente útil, pero puede ser buena.

Segundo, incluso si los esfuerzos de un filósofo están dirigidos a asistir a la racionalidad en la práctica política, pueden útilmente proveer estructuras que hagan visible cuántos conflictos de valor pueden y deben absorberse. En particular, tengo en mente aquí el estudio filosófico de la teoría de las decisiones, que enfatiza estructuras más complejas que los modelos lineales, frecuentemente utilitaristas y centrados en la comparabilidad directa de los valores, que regularmente han prevalecido. Se admite que las estructuras de decisiones reales en los cuerpos políticos se hallan muy lejos de estas idealizaciones, pero también es probable que los modelos simples de la racionalidad práctica puedan afectar los procesos de decisión y ser alimentados por ellos. La filosofía política necesita proveer la combinación de un dominio de las habilidades formales apropiadas y un sentido firme de la existencia de bienes genuinamente competitivos y sólo parcialmente reconciliables —algo que distingue la obra de Amartya Sen en este campo.

Historia

No voy a decir nada, excepto por implicación, acerca del genuino entendimiento histórico. Supongo, por razones obvias y que sería tedioso repetir, que el genuino entendimiento histórico es de la primera importancia para entender la política y el pensamiento político; que es muy importante el distanciamiento diacrónico respecto a nuestra propia sociedad, y todavía más cuando la sociedad que estudiamos es un ancestro de la nuestra. Más aún, me parece que leer a Platón como si fuera un artículo del *Mind* de la semana pasada es perder una parte importante de leer a Platón.

En cuanto a lo que es la comprensión histórica genuina de un texto, el entendimiento de lo que *significa*, estoy de acuerdo con Quentin Skinner en que, si es recuperable, debe serlo del modo en que él detalló, de acuerdo con las expectativas contemporáneas en términos de las cuales se pudiera realizar una intención de comunicar. Entonces, no hay duda de que el significado fundamental de la pregunta formulada acerca de un texto histórico, “¿qué significa?” es “¿qué significó?”

Sin embargo, hay otro tipo de preguntas como “¿qué significa para mí?” o “¿qué es

lo que saco de esto?” Esta pregunta, diría más bien, no es planteada tanto acerca *del* texto, sino *alrededor* de él; ya sea de un libro o un conjunto de palabras que se parecen lo más posible a otras palabras escritas en el pasado. Estas cuestiones, y las respuestas que les damos, me parece que están más bien poco conectadas con la historia, en el sentido del pasado que dio lugar al libro, pero tienen mucho que ver con la *historia del libro*; en particular, la más interesante de estas preguntas puede continuar esa historia. La del libro *Nachleben* no solamente se estudia, sino que se vive.

Cuando tratamos los libros de esta manera y, por ejemplo, si representamos sus argumentos en nuestros propios términos, planteamos preguntas modernas como qué suponemos que son estas preguntas, etc., estamos haciendo algo que es, desde luego, convencionalmente llamado “historia de la filosofía”, pero es realmente un tipo de filosofía —podríamos llamarla “historia de la filosofía de la historia de la filosofía”—. En particular, debemos cuidarnos de dos impresiones engañosas que pueden ser creadas al llamarla “historia de la filosofía”. Una es que la búsqueda de la interpretación lo más consistente que sea posible —una regla básica de este tipo de ejercicio— es un principio genuino de reconstrucción histórica. Podemos pensar que eso se debe, por ejemplo, a su semejanza con un principio que desempeña un papel genuino en el tipo de teoría lingüística que discutí anteriormente: el principio de caridad, por el cual se interpreta un lenguaje extraño de un modo en el cual se concede que lo que es evidentemente verdadero a los hablantes nativos sea efectivamente verdadero, y que no simplemente se estén contradiciendo. Pero ese principio, en ese nivel de generalidad, tiene un papel diferente: no es sólo una suposición contingente, pues no tenemos un control independiente sobre la idea de que, en general, podían estar evidentemente en el error o contradiciéndose (realmente). En el caso de un autor particular y un texto particular, sin embargo, hay en principio más control sobre esta idea, y se pueden alcanzar *a priori* algunas conclusiones acerca de ella. Así, las personas pasan mucho tiempo (yo mismo lo he pasado) tratando de encontrar interpretaciones del *sofista* de Platón que hagan que las teorías de dicho filósofo sean consistentes, pero si el *sofista* trata de lo que pensamos que trata (y concedido que sus teorías acerca de estos temas muy difíciles llegaron cuando las hizo), es muy poco probable que sus teorías sobre esos temas fueran consistentes. Por supuesto, el rechazo del principio de la interpretación consistente deja la entera cuestión en un estado muy aburrido, porque entonces no hay soluciones únicas y hay un número indeterminado de modos en los que podría ser inconsistente. Pero aceptar el aburrimiento como una crítica —de respuestas, quiero decir, no de preguntas— sólo muestra que uno está haciendo filosofía, y en todo caso, que no está haciendo historia.

La segunda consecuencia engañosa de tomar la historia de la filosofía, en el sentido usual, como historia, es que pensamos que tenemos un motivo para hacerlo, por ejemplo, el de proveernos del trasfondo histórico de nuestras propias ideas, pero si el estudio no es genuinamente histórico, entonces no nos provee del trasfondo histórico de nada. Estos textos antiguos tienen una relación en buena medida indeterminada con nuestro entendimiento de lo que significan; son objetos complejos pero ambiguos sobre

los cuales proyectamos conjuntos de ideas filosóficas más que diferentes de las que estaríamos ejercitando en nuestra propia persona. Puede ser interesante, útil, instructivo, incluso de alguna manera enseñarnos algo acerca del pasado, pero nuestra justificación para hacerlo, si es que hay alguna, sólo puede relacionarse con suma complejidad con el hecho de que estos hombres dijeron lo que dijeron con ciertos significados en el pasado. La historia de la filosofía, y en particular la historia de la filosofía política, puede, en principio, ser convertida en historia, pero, como más frecuentemente se hace, especialmente en el espíritu de la filosofía analítica, debe ser defendida, si es defendible, como un tipo de filosofía gracioso, con elementos arcaizantes (algo así como en el estilo del *Pulcinella* de Stravinski), más que como una historia irresponsable.

Conclusión

He tocado ampliamente un número de temas. He dejado implícito que el filósofo político debería estar en contacto con la filosofía moral, y he dicho que, libre de las preocupaciones más estrechas de la distinción hecho-valor, tendrá la oportunidad y la necesidad de estudiar nociones de valor complejas y sustanciales, no sólo en su interrelación conceptual, concebida estrechamente, sino en el trasfondo de las creencias y concepciones no morales que les dan sentido, y que en esto su trabajo debe de hecho interrelacionarse estrechamente con el de las ciencias sociales. He sugerido que dicho trabajo puede requerir ahora un tipo particular de sensibilidad reflexiva que probablemente sea asistida por un estudio de las cuestiones epistemológicas relacionadas con las ciencias sociales, y se puede agregar que también con las ciencias naturales. Debe tener un sentido de las exigencias sistemáticas de la filosofía sin exigir un sistema dentro de las filosofías moral o política mismas, y debe ayudar en algunos aspectos a mantener vivo el sentido de los conflictos genuinamente morales, los orígenes de los cuales muy bien podemos entender, pero que son propensos de ser prematuramente eliminados por medio de la racionalización. Un área en la que se puede hacer esto con utilidad es el desarrollo de estructuras más complejas y realistas en la teoría de las decisiones.

Esto deja en pie la cuestión que molestó tanto a muchos filósofos analíticos de la distinción hecho-valor: ¿Con qué derecho puede un filósofo exigir la atención de un público sobre estos temas, especialmente si —como, ciertamente, he tomado por sentado— sus preocupaciones son en parte normativas? Me parece un signo alentador de lo lejos que ha llegado la filosofía en los últimos quince o incluso diez años, que esta pregunta, que parecía tan honestamente apremiante en un tiempo, parezca ahora tan aburrida, pues su respuesta claramente es que tiene la pretensión que cualquier adulto y persona reflexiva puede tener sobre la atención de otros si tiene pensamientos acerca de algún tema importante. Puede sacrificar esa pretensión, o dejar de hacer lo prometido, en tantos modos como hay palabras escritas muertas, inimaginativas, estúpidas, mal informadas, etc. Puesto que es un filósofo, su llamada a la atención será, como siempre, sopesada con la tendencia hacia un sutil análisis, y a comprobar que algunas de sus

aseveraciones conectan con las de otros y no, por ejemplo, si hay apoyo en amplias y ricas experiencias de la gente o en eventos; pero suponer que eso es todo, y que nada más es eso lo que se le pide, es una tontería desde todos los puntos de vista.

Al decir eso, y en todos los casos, aún tengo en mente a alguien que está haciendo filosofía analítica. He hablado críticamente de algunas limitaciones tempranas de ésta, pero, como dije al principio, no pienso que sus reformas la hayan cambiado o la vayan a cambiar hasta dejarla irreconocible. No tendría caso y sería inútil ahora, no más que al principio, decir cuáles son las que tomo como sus características definitorias, si es que, desde luego, las tiene. La entiendo en una acepción amplia, pero ciertamente mucho de lo que es llamado filosofía queda excluido de ella. Es alentador que se haya vuelto de nuevo algo interesante en el campo de la filosofía política, con sus indudables virtudes intactas. En su insistencia en los valores de la afirmación precisa y el argumento reconocible; su paciencia, su carencia de desdeño a lo familiar, su voluntad de encontrarse con las ciencias formales y naturales; su capacidad para el progreso genuino y discutible, en todo esto, y a pesar de sus muchas y frecuentemente catalogadas limitaciones, sigue siendo la única filosofía real que hay.

XV. La filosofía y el entendimiento de la ignorancia

Nuestro tema es aquello que no sabemos, y ésta es una reunión bajo los auspicios de la filosofía —una reunión, desde luego, diseñada para hacer avanzar el programa en filosofía de la UNESCO—. La conjunción de este tema y estos auspicios nos confronta ya con ciertas preguntas. ¿Cómo está la filosofía relacionada con la ignorancia?

No voy a indagar si la filosofía misma es simplemente un tipo de ignorancia, ni si hay algo así como conocimiento filosófico. Se ha dicho frecuentemente, en particular por los positivistas, que la filosofía es virtualmente, por definición, el hogar de la ignorancia, que comprende preguntas que no sabemos cómo responder mediante formas de investigación establecidas. Conforme a esta explicación, preguntas que han sido previamente parte de la filosofía pueden madurar hasta convertirse en preguntas para las ciencias o para otras disciplinas capaces de acumular conocimiento. De esta manera, algunas preguntas que pertenecieron a la filosofía han pasado a formar parte de la física, la lingüística, la psicología. En la actualidad, varias preguntas sobre la mente y el cuerpo, preocupaciones perpetuas de la filosofía, están (alguien diría) moviéndose al ámbito de la ciencia cognitiva.

El proceso presentado en esta imagen tiene que ser distinguido de otra cosa, a saber, el hecho de que, al menos en el idioma inglés, la *palabra* “filosofía” está más restringida de lo que solía estar. El gran libro de Newton fue una contribución a un tema entonces llamado “filosofía natural”, pero quienes lo practicaban bajo ese nombre eran capaces de trazar distinciones entre él y el tipo de investigación que ahora llamaríamos “filosofía”, y trazarla a lo largo de líneas que nos son, a grandes rasgos, familiares hoy. No es esta cuestión verbal lo que tengo en mente, sino más bien los procesos por los cuales se desarrollan las preguntas a partir de un estatus en el que no sabemos cómo responderlas, y pertenecen a lo que es indisputablemente filosofía, a un estatus en el que se han convertido en un tema propio para la investigación sistemática, y quizá científica. La imagen que enfatiza estos procesos podría, entonces, sugerir no solamente que la filosofía no contiene conocimiento, sino que es virtualmente definida por no tenerlo. Como dije, ésta no es una cuestión que quiera llevar mucho más allá, puesto que en general creo que la filosofía no es de lo más interesante cuando habla acerca de sí misma. Pero quizá pueda decir, al abandonar ese tema, que hay cierta paradoja en pensar que lo que estos procesos demuestran es que la filosofía no ofrece conocimiento de ningún tipo. Ciertamente, hay desarrollos en los cuales —para decirlo muy burdamente— las cuestiones se mueven de la filosofía hacia otros campos de investigación, pero estos

desarrollos no dejan a la filosofía enteramente pasiva en relación con ellos. No es simplemente la cava de vinos en la que cada cuestión madura hasta que alcanza el estado en que puede ser puesta en el mercado de la ciencia, ni es un lugar en el que las cuestiones esperan hasta que se hayan desarrollado otras ciencias, mediante sus propios procesos, hasta un punto en el que sean capaces de recogerlas. La filosofía misma contribuye a estos procesos y, desde luego, puede contribuir tanto que es muy artificial decir que precisamente la misma cuestión se ha graduado de la filosofía a la ciencia. Las cuestiones son reformuladas y redefinidas bajo tales procesos, y aquellos refinamientos y redefiniciones, cualquiera que sea precisamente su estatus, son de hecho un producto de la filosofía.

Lo que deberíamos considerar por más tiempo no es el tema de la filosofía como ignorancia, sino lo que, más bien, la filosofía podría concebiblemente decirnos de un modo más general acerca de la ignorancia. Éste es un tema en el que la filosofía ha mostrado durante siglos un alto grado de ambición: desde el mundo antiguo, a través de la filosofía moderna temprana, hasta los estudios contemporáneos, los filósofos han sido entusiastas al sugerir que sabemos poco o nada. No quiero negar la importancia filosófica de estos argumentos escépticos, pero me parece que dicha importancia yace mucho más en lo que ellos pueden ser capaces de decirnos acerca de la naturaleza o la base del conocimiento más que en una determinación efectiva de qué es lo que no sabemos. La razón de esto es su extrema generalidad: es típico que traten de mostrarnos que no sabemos nada, o que conocemos solamente el contenido de la experiencia inmediata, o que conocemos solamente algunas verdades necesarias simples. Tales conclusiones no hacen que la línea divisoria entre lo que es conocido y lo que no es conocido sea, en ningún sentido, interesante, en particular porque no la representan como una línea que pudiéramos en principio mover. Así que, mientras la posibilidad del escepticismo permanece como un desafío con el que tiene que tratar una teoría del conocimiento de un modo u otro —aunque sea solamente a la manera de algunas teorías contemporáneas, dándole la espalda—, no arroja, por su misma naturaleza, ninguna luz interesante sobre la pregunta de lo que sabemos en oposición a lo que no sabemos.

Esto es quizá solamente un ejemplo de una cuestión más general; que el escepticismo es más interesante y más inquietante si está basado en consideraciones más particulares. Lo que frecuentemente se llama escepticismo “acerca de las otras mentes” es, una vez más, una área importante de la reflexión filosófica, pero cuando se conduce en un nivel de generalidad muy alto, no nos da ninguna razón para ocuparnos de él. Alguien que esté genuinamente preocupado por saber si otra persona tiene dolor, incluso si esa persona se está retorciendo en el piso con un puñal en la pierna, es alguien que debiera ser remitido para tratamiento clínico. Este tipo serio de escepticismo acerca del mundo externo, como el mismo Descartes insistiera, debería reservarse para el ejercicio filosófico reflexivo, pero no es el único tipo de escepticismo acerca de otras mentes, e insistir en el escepticismo en un nivel de generalidad tan alto, realmente sirve para disfrazar la inquietante fuerza de las manifestaciones más particulares del escepticismo sobre las otras mentes. Mientras que es absurdo, en un nivel práctico más que puramente teórico,

preguntarse si otras personas tienen emociones, sentimientos o sensaciones, no es de ningún modo absurdo abrigar dudas reales acerca del carácter de sus sentimientos, o preguntarse cuánto sabemos o podremos saber acerca de la vida interna de otra persona una vez que se va más allá de las características familiares de ella, que son, por virtud del entendimiento humano mutuo, genuinamente dadas a nosotros.

El mismo punto se aplica al escepticismo acerca de nuestro conocimiento del pasado, y también al escepticismo acerca del mundo físico. Plantear la pregunta de si *en absoluto* sabemos algo sobre el pasado es desde luego patológico, si pretende ser más que un artificio para investigar tal conocimiento. Pero una vez más, insistir simplemente en este tipo de escepticismo, y por ende, como reacción a ello, en un rechazo directo del mismo, es oscurecer la extensión y profundidad del escepticismo —que puede estar enteramente justificado— que se puede experimentar con respecto a las narrativas históricas. Hay, desde luego, un problema profundo y estructural acerca de la credibilidad de tales narrativas, y uno deja de aprehender la fuerza de ese problema si plantea solamente la cuestión absolutamente general del escepticismo acerca del pasado, o si, como sucede frecuentemente, se extiende el escepticismo apropiado a la narrativa histórica incluso a los elementos menos ambiciosamente reportados de una crónica. Es absolutamente cierto que hubo un sismo en Los Ángeles en enero de 1994, y que Julio César fue asesinado en Roma en marzo del 44 a.C., y solamente si tal nivel de afirmación acerca del pasado es aceptable, uno puede proceder a plantear las preguntas realmente interesantes de qué tanto ignoramos acerca del pasado.

Nuestra preocupación sobre lo que no sabemos puede apoderarse de nosotros en forma abrumadora solamente si hay ciertas cosas que sí sabemos. Esta es la razón por la cual la filosofía escéptica tradicional, al sugerir —y de hecho lo hace— que no podemos saber absolutamente nada (nada sobre otras mentes, nada sobre el pasado), no es convincente, al menos en el sentido de atacar nuestra seguridad sobre estas cosas. Aún más, no se nos debe convencer —y no se puede— con reflexiones muy generales dirigidas no sólo contra el conocimiento sino contra la verdad. Formas recientes de escepticismo, que se apoyan en muchos casos en una lectura muy parcial de Nietzsche, han tratado de desacreditar totalmente la noción de verdad. Al hacer esto, suelen adoptar un tono de suave heroísmo por su proyecto de poner en evidencia nuestras ilusiones (como a ellos, desviadamente, les gusta decirlo).

Lo inquietante de tales posiciones no es tanto que sean contradictorias sino sus falsas promesas de incomodar. Lo que quiere hacernos sospechar de todo no nos hace sospechar de nada; incluso el paranoico común o de jardín necesita su libro de ejercicios de hechos cuidadosamente investigados. Nuestras seguridades sospechosas serían socavadas, como lo son en la propia práctica de Nietzsche, solamente por una atención interpretativa que se dirija selectivamente, y que acepte los materiales necesarios para que pueda entenderse hacia donde está dirigiendo su atención. En el caso de la historia, podemos tener dudas acerca de nuestro entendimiento del pasado sólo si tenemos un pasado, y tenemos un pasado (como enfatizaba Wittgenstein) sólo si hay algunas cosas que con certeza ocurrieron en el mismo. Debido a que podemos aceptar un gran número

de hechos acerca del pasado, muchos de ellos en sí mismos muy aburridos, podemos confrontar la sugerencia genuinamente perturbadora de que el entendimiento histórico requiere narrativa, y la narrativa demanda que se dejen de lado algunos acontecimientos, lo que siempre genera en la historia una ficción y frecuentemente una mentira.

En el caso del mundo físico como tal, la línea divisoria entre las versiones convincentes y las no convincentes del escepticismo adopta una forma diferente. Un escepticismo total acerca del mundo externo no es convincente, en el sentido de que le dé a uno alguna razón para dudar si todo lo que pensamos que sabemos acerca del mundo externo podría ser falso. Igualmente, puede no haber razones para suponer que no sepamos un número muy grande de verdades basadas en la observación, asistida o no por aparatos científicos. La pregunta que aquí nos interesa surge con respecto a nuestras proposiciones teóricas acerca de la naturaleza, y las entidades teóricas que son normalmente introducidas por éstas. Es todavía una cuestión muy viva en la filosofía de la ciencia en qué medida nuestro entendimiento teórico de la naturaleza es conocimiento, pero hay dos niveles diferentes en los cuales surge tal preocupación. Por un lado, está la pregunta que frecuentemente es expresada como una elección entre el instrumentalismo y el realismo, y que nos plantea si *cualquier* teoría podría ser conocimiento, o si nuestro acceso real al conocimiento de la naturaleza se halla confinado simplemente al nivel de lo que es observable (donde lo que es “observable” involucra por sí mismo enormes dificultades). Por otro lado, podemos preguntar si tales teorías *y no otras*, son conocimiento. En este nivel, ya no se pregunta si hay algo inherente en tales teorías (en tanto que opuestas a las observaciones) que les impida ser conocimiento, sino, más bien, si algunas teorías científicas tienen mejores bases para pretender ese estatus que otras.

Se puede pensar que esta última pregunta no es en sí misma filosófica, sino más bien, precisamente, el tipo de pregunta que forma la sustancia de la práctica científica, la cual se preocupa por avanzar y preferir algunas teorías a otras. Sin embargo, pienso que es un error asociar siempre el término “filosófico” con los tipos más generales de pregunta que pueden surgir con respecto a las teorías o las hipótesis. Dicho de otra manera, hay áreas del saber en que la línea divisoria entre la filosofía de la ciencia y la ciencia misma es trazada muy arbitrariamente. Un ejemplo muy claro de esto es la mecánica cuántica, en la cual es fundamental aceptar los límites conceptuales requeridos por las descripciones y explicaciones que ofrece. Mientras que nadie puede discutir estos asuntos sin estar bien informado acerca de la situación actual de tal teoría física, más allá de ese punto se vuelve muy arbitrario, en muchos casos, determinar si alguien está discutiendo la teoría filosófica de la mecánica cuántica, o haciendo mecánica cuántica en un nivel altamente teórico.

Algunas de las cuestiones filosóficas que he mencionado son irresolubles por el momento, pero podemos tener esperanzas, si no de resolverlas, al menos de avanzar en nuestro entendimiento de ellas en los años venideros. Me parece que esto es verdadero con respecto a ciertas cuestiones que acabo de mencionar: la relación entre observación y teoría en la filosofía de la ciencia, y entre la estructura de la narrativa histórica y lo que he llamado los materiales de la crónica. Se puede decir razonablemente que en el

presente no sabemos cómo analizar de la mejor manera estos temas, y no me parece tonto suponer que podamos llegar a entenderlos mejor en el futuro. Estos asuntos son propios de la filosofía, de modo que si podemos tener esperanzas de progresar en ellos, eso sería un ejemplo de que existen preguntas *dentro* de la filosofía que fueron contestadas más satisfactoriamente que como lo habían sido hasta entonces. Por esa misma razón, no deberíamos, como ya he sugerido, prestar demasiada atención a los argumentos que quieren mostrar que nunca podremos llegar a saber nada en filosofía. La frase *philosophia perennis*, podríamos decir, expresa pesimismo. Ni es la relevancia de la filosofía nuestro tema, ni ha de encontrarse principalmente en su vieja invitación a suponer que no sabemos nada en absoluto.

Hay, sin embargo, un argumento filosófico muy diferente que incide en nuestro tema, y que pienso que es más interesante para el mismo. Sugiere que, mientras que sin duda podemos formular algunas cosas que no sabemos, no podemos en general formular con confianza qué es lo que no sabemos. Hay límites de principio a la medida en que podemos saber lo que no sabemos. El punto se puede demostrar, de una manera más bien restringida, mediante el siguiente argumento.[111] Supóngase que estoy mirando una habitación con muchas personas en ella. No sé cuántas son y no estoy en posición de contarlas, pero sé (por ejemplo) que hay menos de mil personas, o de modo general, para varios números n , sé que hay menos de n personas. Cuando n se vuelve más pequeño (más cercano a lo que es, de hecho, el número real de personas), deja de ser cierto que sepa que hay menos de n personas presentes. Sin embargo, es obvio que no hay un número particular n tal que sepa que hay menos de n personas presentes pero no si hay menos de $(n - 1)$: así que, para algunos números, debe ser cierto que no sé si no sé que hay menos de ese número de personas presentes.

Ahora bien, esto es, desde luego, un ejemplo extremadamente artificial y restringido, pero ilustra un punto mucho más general acerca del conocimiento de la ignorancia. En general, hay límites importantes al conocimiento de la ignorancia, y este punto se halla implícito en la noción de margen de error, una noción que está en sí misma relacionada con el concepto de conocimiento. No sé, en todo caso, lo que no sé; también es verdad que, aunque no sepa algo, no necesariamente sé que no lo sé: el principio tan amado de ciertos filósofos, de que si sé, debo saber que sé (el así llamado “principio KK”), es ciertamente falso. Esto tiene consecuencias importantes para cualquier investigación del tipo de la que estamos abordando. Desde luego, podemos ser capaces de mencionar algunas cosas que no sabemos; ya he mencionado una o dos de ellas, y procederé a mencionar una o dos más. En algunos casos, es posible especificar qué es lo que uno no sabe de tal modo que uno puede saber lo que uno no sabe: así, en el ejemplo previo, sé que no sé *el número exacto de personas que hay en la habitación*. De modo similar, yo personalmente expreso conocimiento si digo que no sé el nombre del actual arzobispo de Milán, o del padre de Leonardo da Vinci. Pero no siempre sabemos que no sabemos ciertas cosas, y, como ya se ha argumentado, se puede mostrar que es imposible saber en todos los casos que no sepamos algo.

Hay otros argumentos menos formales que conducen a la misma conclusión. Entre

las cosas que no sabemos, hay algunas que las personas solían saber, y otras que las personas podrán llegar a saber en el futuro. Hay algunas asimetrías importantes entre estas dos clases. Una de estas asimetrías es que, con respecto a las cosas que las personas solían saber pero que nosotros no sabemos, suponemos en general que poseemos los términos en los que tal conocimiento podría expresarse. Conocemos o sabemos de los términos en que la gente del pasado hubiera expresado su conocimiento; términos que podríamos ser capaces de entender en algún nivel. Esto es porque nuestra imagen del conocimiento que se ha perdido es expresada (muy a grandes rasgos) en términos que no se han perdido.

Sin embargo, respecto del conocimiento que las personas pueden adquirir en el futuro, no necesariamente tenemos esa concepción. Fue un error positivista, al cual nadie se apega ahora, suponer que el vocabulario fundamental o los recursos conceptuales de la ciencia son fijos, y que lo que será descubierto en el futuro podría necesariamente ser expresable en ese mismo vocabulario. Por el contrario, creemos que los avances teóricos implican por lo regular la introducción de nuevos conceptos, que pueden no ser estrictamente compatibles con los que actualmente tenemos. No pienso que esto tenga que conducir a un relativismo radical, pero significa que en el futuro la ciencia podría contener innovaciones teóricas que, como están las cosas, no entenderíamos en absoluto. Es un tema de discusión si podría llegar a haber innovaciones tales que no pudiéramos en principio llegar a entender. Quizá, como algunos han sugerido, nuestra idea de un lenguaje posible debe ser la idea de un lenguaje que podríamos llegar a entender en principio, pero no necesitamos inmiscuirnos en ese asunto; nuestra cuestión presente tiene que ver solamente con la posibilidad de que los futuros descubrimientos sólo puedan expresarse en un lenguaje que ahora, de hecho, no podríamos entender, y eso ciertamente no puede ser eliminado por ningún argumento de principio. Tales descubrimientos futuros, suponemos, serían *descubrimientos*, lo que quiere decir que podrían constituir conocimiento, pero no podemos saber qué será ese conocimiento, por la radical razón de que no tenemos manera de expresarlo; consecuentemente no podemos saber qué es exactamente lo que, al carecer de ese conocimiento, no sabemos. Podemos ubicarlo en varios grados, y debemos hacerlo si hemos de considerarlo un descubrimiento científico posible —debemos ubicarlo en un campo de estudio, por ejemplo—, pero eso está muy lejos de permitirnos saber con exactitud qué es lo que no sabemos.

La consideración de que son previsibles los términos en que podrían expresarse los descubrimientos futuros, se aplica particularmente a ciertas áreas de la investigación científica que actualmente nos hacen pensar que hay no solamente mucho que no sabemos —esto es obviamente verdad en cada área—, sino que podría haber una cierta inseguridad respecto del conocimiento que esperamos ya poseer. No pretendo ser ningún experto en física teórica, y lo que tengo que decir acerca de ese tema es de lo menos profesional, pero la opinión experta de otros que se hallan mejor ubicados para entender la situación actual puede conducirme a pensar que, con respecto tanto a la cosmología científica como a la física de partículas (campos estrechamente relacionados), las estructuras de la teoría que tenemos se hallan en más de un modo peligrosamente

relacionadas con lo que definitivamente sabemos. En el caso de la cosmología, las conclusiones o especulaciones actuales pueden estar, más bien, corroboradas en forma extensiva con observaciones e interpretaciones de éstas a tal grado que la alteración de una suposición acerca del principio del argumento nos puede llevar a reajustes radicales de todo el panorama teórico. Entre estas conclusiones, quizá, se hallan los órdenes de magnitud del tiempo y la distancia que se invocan en la teoría, y puede ser razonable pensar que en estos y otros aspectos pudiera estar esperándonos una revisión radical de la teoría actual.

En este punto, la situación es diferente en otras áreas igual de teóricas de la ciencia (lo que ilustra el punto que mencioné antes, de que las distinciones filosóficas entre teoría y observación no coinciden con las preguntas que nos ocupan aquí). Pienso que hay un acuerdo general en que es indudable que nuestro entendimiento de la estructura del ADN y la mecánica de la transmisión de la información genética son fundamentalmente correctos. Nuestro entendimiento de esto parece ahora haber alcanzado el punto en que mucho de él formará parte de los datos de cualquier teoría futura, y no de lo que podría ser reemplazado por ella.

El caso de la física de partículas, en la medida (nuevamente, muy limitada) en que la entiendo, es en algunas maneras similar al de la cosmología, pero aquí, incluso quienes confían en que la teoría actual es básicamente correcta, son conscientes de la limitación al desarrollo del conocimiento que yace en los límites económicos y prácticos de la experimentación controlada. Supongo que es posible que el súper colisionador europeo, si, desde luego, finalmente es fabricado, pudiera ser la última máquina de su especie jamás construida sobre la tierra, y sabemos ya que no toda pregunta que pueda plantearse dentro de la estructura presente de la física de partículas, incluso suponiendo que esa estructura continúe sin revisión radical, puede ser respondida por esa máquina. Ésta es una área en que la teoría ha rebasado con mucho cualquier experimento previsible. Si esto es así, entonces nuestro conocimiento de nuestra ignorancia tendría una estructura especial; reflejaría el hecho de que, incluso si se plantean bien las cuestiones, la cantidad de energía (y, por ende, de recursos) que sería necesaria para conducir experimentos que pudieran responderlas simplemente rebasaría cualquier posibilidad real. Esas cantidades de energía se hallan, desde luego, en operación en el universo, y los laboratorios finales de la física de partículas son las estrellas, pero no podemos conducir experimentos controlados con las estrellas, y la brecha entre lo que podemos sensiblemente construir para los propósitos del experimento, y las energías requeridas por los experimentos que pudieran responder a nuestras preguntas, pueden ser por siempre imposibles de cerrar.

Estas especulaciones acerca de las ciencias tienen, como lo he dicho, muy poca autoridad. Sin embargo, pueden al menos ilustrar un punto general que podría ser pasado por alto por la filosofía, a saber, que es posible lograr desde dentro de la ciencia alguna idea de las preguntas que es posible o no responder, incluso si concedemos que los términos en que una respuesta pudiera ser dada pueden variar en grados que nos son desconocidos. Esta aparente paradoja se puede resolver porque en alguna medida

podemos formarnos una idea de la situación experimental en la cual tendría que buscarse una respuesta, si es que tal cosa fuera posible, y podemos ver que somos incapaces de llegar a esa situación, o que es improbable que alguna vez seamos capaces de hacerlo. Todo esto debe ser, inevitablemente, especulativo, porque, al no prever la teoría futura, tampoco podemos prever el experimento futuro; quizá se pueda llegar a disponer de maneras menos caras, y que actualmente no podemos concebir, de investigar la naturaleza en la escala micro. Ciertamente, en estas áreas nuestros pensamientos acerca de lo que no sabemos deben tomar en consideración las rutas experimentales existentes que podrían volverse disponibles para nosotros, y también, como es notablemente el caso de la cosmología, la pregunta por el grado en que las partes más ambiciosas de la teoría se derivan del material indiscutiblemente observado. En esa medida, hay una posibilidad de que, bajo la revisión de algún elemento de la estructura teórica, ésta se derrumbe y cambie a algo muy diferente. Esto nuevamente ilustra el punto de que es un error que los filósofos simplemente discutan el estatus de la ciencia teórica *como tal*. La física de partículas, la cosmología y, desde luego, la biología molecular, son todas ciencias teóricas, pero sus relaciones con el experimento y la refutación futuras, y correspondientemente con nuestra ignorancia presente, son muy diferentes entre ellas.

Hay un ejemplo drástico de nuestra ignorancia actual, cuyas relaciones tanto con la filosofía como con la investigación científica son muy diferentes de lo que se ha discutido hasta aquí. La principal diferencia es que en este caso sí es una pregunta filosófica hasta qué punto nuestra ignorancia tiene algo que ver con la investigación científica. Hablo del problema de la conciencia, la cuestión de cómo podemos explicar, o incluso describir adecuadamente, la diferencia entre criaturas *para* las cuales las cosas existen u ocurren de ciertos modos, y las criaturas que carecen de tal experiencia. Mucho trabajo se ha hecho en esta área en años recientes, y se está haciendo ahora, y al menos se puede decir que poseemos más material del que teníamos antes, sobre todo en el nivel neurofisiológico, que pudiera resultar relevante para una solución de este problema, si es que ésta es posible. Mi formulación extremadamente cauta de este hecho está dictada por el muy amplio rango de opiniones que es aún posible mantener sobre este asunto. En un extremo se hallan personas, tales como Daniel Dennett, que sostendrían que al menos en bosquejo han resuelto el problema. Y luego están aquellos, quizá la mayoría, que piensan que el problema no ha sido resuelto, pero que alguna combinación de investigación fisiológica, trabajo en ciencia cognitiva y clarificación filosófica, debería o podría eventualmente darnos la solución. Nuevamente, se hallan quienes, como Thomas Nagel, creen que por el momento no solamente no tenemos manera de relacionar los hechos de la conciencia, las experiencias de primera persona de un agente consciente, con cambios en el cerebro y el sistema nervioso, sino que incluso carecemos de cualquier idea de qué aspecto tendría una teoría que relacionara aquellas dos cosas de una manera clara y explicativa. Alguien que piensa esto puede, no obstante, suponer que esta situación puede ser como la que ha precedido a otros avances científicos, en los que un paso que ha parecido incomprendible para nosotros ha sido, de hecho, dado mediante una nueva invención conceptual.

En el más remoto extremo del pesimismo sobre el tema de la conciencia se halla Colin McGinn, quien sospecha que el problema de dar una descripción coherente y explicativa de la conciencia humana es irresoluble por una razón muy especial, a saber, que la estructura del cerebro es tal que tal vez no pueda aprehender este aspecto de sus propias operaciones. Esta conclusión, naturalmente, y quizá sanamente, atrae al escepticismo, porque suena mucho como aseveraciones previas acerca de lo que pudiera resultar científicamente ininteligible, en particular el así llamado problema de la naturaleza de la vida, que dio lugar a la posición pesimista y oscurantista del vitalismo. Sin embargo, esta analogía no necesariamente nos conduce a rechazar el pesimismo de McGinn (que, en cualquier caso, está mejor argumentado de lo que jamás lo estuvo cualquier posición vitalista), sino que nos lleva también a un pensamiento adicional acerca de lo que podría suceder con este problema. En cierto sentido, el problema de la naturaleza de la vida —el problema que el vitalismo declaró irresoluble— nunca fue del todo resuelto: más bien, hemos aprendido tanto acerca de las operaciones de los seres vivos que el problema, en esa forma, ha desaparecido. Como están las cosas en la actualidad parece que el problema de la naturaleza de la conciencia podría no ser éste, puesto que la conciencia parece un fenómeno tan presente y manifiesto que no podemos entender cómo es que la cuestión de su naturaleza podría, por virtud de una enorme elaboración de entendimiento fisiológico y psicológico, evaporarse en última instancia. Pero quizá esa impresión misma es una función del estado presente de nuestro entendimiento o, más bien, de nuestra carencia de él.

Ciertamente, el problema de la conciencia combina en el más alto grado los diversos tipos de duda que constituyen la admisión de nuestra propia ignorancia. No estamos de acuerdo en que hay un problema, o si lo hay, en que haya sido resuelto, o si no lo ha sido, en que sea resoluble, o si es resoluble, en que los obstáculos que presenta para su resolución sean técnicos, teóricos o conceptuales. Por añadidura, no estamos de acuerdo en que el problema, si no puede ser directamente resuelto, podría desaparecer en un momento dado. Es difícil pensar en otra área de problemas en la cual tantas impresiones de la naturaleza de nuestra ignorancia puedan coexistir.

Hay al menos un problema que, particularmente, quizá mientras discutimos estos asuntos bajo los auspicios de la UNESCO, debería mencionarse, que desafía nuestro entendimiento y revela nuestra ignorancia: el problema de cómo vivir juntos. Nadie va a negar, presumiblemente, que haya tal problema, en el sentido de que varios grupos constante y repetidamente, por todo el mundo, encuentran sumamente difícil vivir al lado de otros. En este sentido, ciertamente, hay algo que no sabemos: cómo vivir juntos —excepto en ciertas circunstancias adecuadamente favorables—. Sin embargo, es una afirmación diferente, y podría pensarse que es más bien una aseveración optimista, que el problema de que no sabemos por qué no sabemos cómo vivir unos con otros es un problema intelectual. Esto sugiere que hay algo que investigar acerca de las causas del conflicto; algo que está actualmente escondido de nosotros y que, cuando sea encontrado, podría permitirnos negociar y eliminar progresivamente esos conflictos. Quizá hay algo así que está actualmente escondido de nosotros, y ciertamente no

deberíamos relajar nuestros esfuerzos preguntando qué podría ser, buscando la ayuda de la psicología, la antropología, la historia y tal vez la biología. Sin embargo, no podemos estar seguros de que exista dicho problema intelectual. Quizá, más bien, ya sepamos más de lo que es sabido a un nivel general acerca de las causas de conflicto humano, y no haya nada muy profundo o extenso, que no reconozcamos ya, de donde podríamos aprender. Lo que necesitamos es, más bien, organizar los recursos que en términos generales sabemos que son necesarios para tratar con tales conflictos, en tanto que pueden ser tratados, entendiendo cada uno según sus propias circunstancias. Si no podemos movilizar los recursos, o no es el tipo de conflicto que podría responder a cualesquiera recursos que pudiéramos movilizar, no supondremos que hay una cosa adicional, potencialmente reveladora, que no sepamos. Tendremos que asumir una cosa perfectamente obvia, que no sabemos cómo tratar con el conflicto.

No tenemos necesariamente que aceptar esta segunda explicación, más sombría. Quizá la idea más feliz, de que hay aún algo importante que aprender acerca del hecho de que los seres humanos estén tan dispuestos a odiarse y matarse entre sí, ofrece alguna promesa. Si la ofrece, ciertamente nos gustaría saber que lo hace. Si la historia más sombría es verdadera, sin embargo, quizá no queramos saber que lo es. En éste, el más apremiante de nuestros asuntos, tomamos la posición que solemos tomar con lo que afecta a nuestros intereses: no podemos saber si realmente queremos superar nuestra ignorancia hasta que lo hayamos hecho.

XVI. La filosofía como una disciplina humanística[j]

1

En la fórmula “disciplina humanística” se pretende que ambos términos carguen un peso. Ésta no es una conferencia acerca de la organización académica: al hablar de filosofía como una empresa “humanística”, no estoy sugiriendo que ésta pertenezca a las humanidades o a las artes. La cuestión es: ¿qué modelos, ideales o analogías debiéramos buscar al pensar en los modos en que debiera practicarse la filosofía? Se trata de una aplicación a nuestras circunstancias presentes de una cuestión más general y tradicional, la cual es notoriamente una cuestión filosófica: ¿cómo debería la filosofía entenderse a sí misma?

Sucede lo mismo con el otro término de la frase. No se trata de disciplina como un campo o área de investigación. Se supone que “disciplina” implica disciplina. En filosofía debe haber algo que cuente para lograrlo como se debe, o hacerlo bien, y yo creo que esto debe todavía estar asociado con las metas de la filosofía de ofrecer argumentos y expresarse con claridad, metas a las que se les ha dado énfasis particularmente en la filosofía analítica, aunque a veces de una manera perversa y unilateral. Pero la expresión clara y la argumentación no son monopolio de la filosofía. Otras materias humanísticas ofrecen argumentos y se pueden expresar claramente, y si no pueden hacerlo, ése es su problema. La historia, por ejemplo, tiene sus disciplinas, las cuales implican, entre otras cosas, tanto argumentos como claridad. Supongo que la historia es central en un estudio humanístico, y no hace ninguna diferencia en esto que la historia, o algunos aspectos de la historia, sean a veces clasificados como ciencia social; eso solamente nos dirá algo acerca de cómo entender la idea de una ciencia social. La historia es central para mi argumento no sólo porque es central entre las disciplinas humanísticas, sino porque, como voy a argumentar, la filosofía tiene relaciones muy especiales con ella.

Una cierta relación limitada entre historia y filosofía ha sido tradicionalmente reconocida, al grado de que se esperaba que las personas que estudiaran algo de filosofía aprendieran también algo de historia de la filosofía. Esta idea tradicional ya no es aceptada en todas partes ahora, y regresaré después a ese punto. Debe decirse, también, que esta tradicional concesión a la historia fue frecuentemente más bien nominal: muchos de los ejercicios conducidos en nombre de la historia de la filosofía han tenido una relación muy tenue con algo que pudiera independientemente ser llamado historia. La actividad fue identificada como “historia de la filosofía” más por los nombres que

aparecían en ella que por los modos en que se conducía. Paul Grice solía decir que “deberíamos tratar a los filósofos grandes y muertos como tratamos a los grandes y vivos, como personas que tienen algo que decirnos a *nosotros*”. Eso está bien, en tanto que no se suponga que lo que los muertos tienen que decirnos es más o menos lo mismo que lo que tienen que decirnos los vivos. Desafortunadamente, esto es probablemente lo que han supuesto quienes, en el clímax de la confianza en lo que ha sido llamado “la historia analítica de la filosofía”, nos motivaban a leer algo escrito por Platón “como si hubiera aparecido en *Mind* el mes pasado” —una idea que, si es que significa algo, es que destruye la razón principal de leer a Platón.[112]

El punto no está confinado al estilo “analítico”. Hay un agradable pasaje de Collingwood en el que describe cómo “la vieja pandilla de los realistas de Oxford”, como él llamaba a Prichard y a Joseph, insistirían en traducir alguna expresión griega antigua, como “obligación moral”, para señalar entonces que Aristóteles, o quien quiera que haya sido, tenía una teoría inadecuada de la obligación moral. Era como una pesadilla, decía Collingwood, en la que se conocía a un hombre que insistía en traducir la palabra griega *trirreme* como “barco de vapor”, y luego se quejaba de que los griegos tenían una concepción defectuosa de los barcos de vapor. Pero, en cualquier caso, lo que quiero decir acerca de la relación de la filosofía con la historia va mucho más allá de su preocupación por su propia historia, aunque ésta es, desde luego, parte de aquélla.

Ya he hablado acerca de que la filosofía es esto o aquello, o que tal cosa y tal otra son centrales para la filosofía, lo que puede ya haber provocado la sospecha de esencialismo, como si la filosofía tuviera alguna naturaleza enteramente exclusiva y atemporal a partir de la cual se pudieran extraer varias conclusiones. Así que permítaseme decir de una vez que no quiero caer en ninguna idea de ese tipo. Desde luego, sostendré posteriormente que algunas de las compenetraciones más profundas de la filosofía moderna, especialmente en la obra de Wittgenstein, permanecen sin desarrollarse —y que, desde luego, yéndose al límite, llegan a ser ininteligibles— precisamente debido a una suposición de que la filosofía es algo muy peculiar, que no debería confundirse con ningún otro tipo de estudio, y que no necesita otro tipo de estudio para entenderse a sí misma. Wittgenstein en su última obra rechazó el esencialismo de un mundo que ha sido muy influyente, y habló de semejanzas de familia y otras cosas por el estilo, pero al mismo tiempo estaba obsesionado —y no pienso que sea una palabra demasiado fuerte— por la identidad de la filosofía como una empresa completamente peculiar; esto es así en la concepción de Wittgenstein, ya sea que se interprete esto como que la compulsión a penetrar en ella es patológica, o parte de la condición humana.[113] No me parece que la filosofía sea así de peculiar, y, por añadidura, debería recordar el punto en que Wittgenstein nos invita a traer a colación otras cosas, lo que es muy diferente. Lo que tengo que decir se aplica, espero, a la mayor parte de lo que es considerado de modo estándar como filosofía, y trataré de explicar por qué es así, pero no trataré de deducirlo de la naturaleza de la filosofía comparada con otras disciplinas, ni de ninguna otra cosa. Lo que tengo que decir, puesto que es ello mismo una pieza de filosofía, es un ejemplo de lo que supongo que es la filosofía: parte

de una intención más general de hacer lo mejor de nuestra vida, y así para nuestras actividades intelectuales, en la situación en que nos encontramos.

2

Un claro contraste definido a una concepción humanística de la filosofía es el *cientificismo*. No quiero decir con esto simplemente un interés o profundización en la ciencia. La filosofía debería estar interesada en las ciencias y algunos filósofos muy bien pueden profundizar en ellas, y nada de lo que digo pretende negarlo. El *cientificismo* es, más bien, un malentendido de las relaciones entre la filosofía y las ciencias naturales que tiende a asimilar la filosofía a las metas, o al menos a las maneras, de las ciencias. De acuerdo con lo que dije recientemente acerca de la variedad de la filosofía, hay ciertamente algún trabajo filosófico que es muy correcto que se conduzca como una extensión de las ciencias naturales o las matemáticas, porque eso es lo que es: la filosofía de la mecánica cuántica, por ejemplo, o los aspectos más técnicos de la lógica, pero en muchas otras áreas la asimilación es un error.

No quiero decir mucho acerca de lo que podría ser llamado “*cientificismo estilístico*”; como la pretensión, por ejemplo, de que la filosofía de la mente es el extremo más teórico y menos experimentalmente encumbrado de la neurofisiología. Puede sugerirse que este tipo de asimilación, incluso si es hasta cierto punto extraviada, al menos alienta un cierto tipo de rigor, que ayudaría a cumplir la promesa de la filosofía de dar cuerpo a una disciplina, pero dudo que esto sea así. Por el contrario: la filosofía *cientificista* de la mente no puede encarnar el rigor que es en primera instancia apropiado a la neurofisiología, el de los procedimientos experimentales; las contribuciones filosóficas de este tipo son las que más probablemente habrán de semejarse a otro fenómeno bien conocido en los círculos científicos, en los que los investigadores improvisan declaraciones en sus periodos de descanso, apropiadas para reuniones informales. Estas declaraciones son muy frecuentemente interesantes, pero ello es porque son observaciones de los científicos, que están dichas por detrás de lo que ordinariamente hacen. No hay muchas razones para esperar tanto interés en las observaciones de los filósofos que no están de vacaciones de nada, sino cuyo trabajo se supone que es hacer tales observaciones.

Una cuestión que me intriga y de la cual no conozco la respuesta es la relación entre una concepción científica de la filosofía y el estilo bien conocido y usual de muchos textos de filosofía analítica que buscan precisión a través del total control de la mente, a través de la emisión de direcciones interpretativas continuas y rígidas. De un modo que será familiar a cualquier lector de filosofía analítica, y es demasiado familiar a todos los que la perpetramos, este estilo trata de eliminar por anticipado todo malentendido, mala interpretación u objeción concebible, incluyendo las que vendrían solamente de los maliciosos o los que clínicamente no pueden entender nada sino de un modo literal. Frecuentemente, esta actividad en sí es justificada, lamentablemente, por la alardeada

necesidad de rigor y claridad de la filosofía analítica. Ahora bien, es perfectamente razonable que un autor deba considerar las objeciones y posibles mal entendidos, o al menos muchos de ellos; lo extraño es que los ponga en un texto. Es de esperarse que todos esos elementos, al ser considerados, le den forma a la obra, pero, como se hace al construir un edificio, una vez que éste queda terminado, el andamiaje se elimina para permitir el libre acceso de la gente.

Hay sin duda más de una motivación de este estilo. Una es la enseñanza de la filosofía mediante discusiones polémicas, la cual tiende a crear en los filósofos un súper ego intimidante y escrupuloso; una mezcla de sus maestros más impresionantes y sus colegas más competitivos, que guía su escritura mediante anticipaciones constantes de culpa y vergüenza. Otra son los requerimientos del doctorado como un ejercicio académico, el cual implica la producción de un tipo de texto muy peculiar, que puede ser muy fácilmente confundido con un libro. Hay exigencias de promoción académica que pueden alentar a uno a publicar tantas páginas como sea posible a partir de cualquier idea que se tenga por modesta que sea. Ahora, ninguna de estas influencias está necesariamente conectada con una visión científicista de la filosofía, y muchas de las personas que optan por este estilo rechazarían con razón cualquier sugerencia de que tuvieran esa perspectiva. Desde luego, un ejemplo obvio de esto es un filósofo que quizá hizo más que cualquier otro para alentar este estilo, G. E. Moore. Sin embargo, por todo ello, pienso que no deberíamos rechazar demasiado rápidamente el pensamiento de que, estando el científicismo donde está, este estilo puede transformarse en espíritu científicista. Puede servir como un remedo del procedimiento científico escrupuloso. Las personas quizá puedan convencerse a sí mismas de que si persisten lo suficiente con evaluaciones y ejemplos comparativos, están ejerciendo el equivalente filosófico de un protocolo bioquímico.

3

Pero como dije, el científicismo estilístico no es realmente la cuestión que nos ocupa; hay un asunto mucho más importante aquí. Consideremos aquí el siguiente pasaje del libro de Hilary Putnam para las conferencias Gifford, titulado *Renewing Philosophy*:^[114]

La filosofía analítica ha estado crecientemente dominada por la idea de que la ciencia, y solamente la ciencia, describe al mundo como es en sí mismo, independiente de las perspectivas particulares. Seguramente, hay dentro de la filosofía analítica figuras importantes que combaten ese científicismo [...] No obstante, la idea de que la ciencia no deja espacio a una empresa filosófica independiente ha alcanzado el punto en que los practicantes líderes a veces sugieren que todo lo que queda para la filosofía es tratar de anticipar qué apariencia tendrán las presuntas soluciones científicas a todos los problemas metafísicos.

No es difícil encontrar aquí un importante *non sequitur*. ¿Por qué la idea de que la ciencia y solamente la ciencia describe al mundo como es en sí mismo, independiente del punto de vista, debería significar que no hay una empresa filosófica independiente? A esta suposición se podría llegar si se asume que, si hay una empresa filosófica

independiente, su meta es describir el mundo como es en sí mismo, independientemente de las distintas perspectivas. Y ¿por qué debiéramos aceptar eso? Admito que resiento ese *non sequitur* porque, a lo largo del libro de Putnam (el cual contiene un capítulo llamado “Bernard Williams y el saber absoluto del mundo”), yo mismo soy identificado como alguien que “ve la física como si nos diera la verdad metafísica última...”[115] Ahora bien, nunca he sostenido tal concepción, y estoy totalmente de acuerdo con Putnam al rechazarla. Sin embargo, he abrigado la idea de que la ciencia podría describir el mundo “como es en sí mismo”, es decir, dar una representación del mismo que sea en la mayor medida posible independiente de las perspectivas particulares o idiosincrasias locales de los investigadores, una representación del mundo, como diría yo, “es de cualquier modo”. [116] Llamé a tal representación en mis términos “el saber absoluto del mundo”. Si es alcanzable o no, o si la aspiración a la misma es incluso coherente, son por supuesto cuestiones muy discutibles.

Un signo de que algo debe haber fallado en el argumento de Putnam, o en el mío, si no es que en ambos, es que este autor supone que la idea de un saber absoluto del mundo debe en última instancia estar motivada por la meta contradictoria e incoherente de describir el mundo sin describirlo: como lo dice él,[117] no podemos dividir el lenguaje en dos partes, “una parte que describe el mundo ‘como de todos modos es’, y una parte que describe nuestra contribución conceptual”. (La siempre tramposa palabra “nuestra” es importante, y habremos de regresar a ella.) Pero mi meta al introducir la noción del saber absoluto era precisamente darle la vuelta al punto de que uno no puede describir el mundo sin describirlo, e introducir la compenetración fundamentalmente kantiana de que simplemente no hay una concepción del mundo que no esté conceptualizada de un modo u otro. Mi idea no era que pueda conceptualizarse el mundo sin conceptos, sino que, cuando reflexionamos sobre nuestra conceptualización del mundo, podríamos ser capaces de reconocer desde dentro de ella que algunos de nuestros conceptos y modos de representar el mundo son más dependientes que otros de nuestra propia perspectiva, nuestras maneras peculiares y locales de aprehender las cosas. En contraste, podríamos ser capaces de identificar algunos conceptos y estilos de representación que sólo son mínimamente dependientes de nuestros modos peculiares, o de los de cualquier otra criatura de aprehender el mundo: éstos formarían una especie de representación que podría ser alcanzada por cualesquiera investigadores competentes del mundo, aunque difirieran de nosotros —es decir, de los seres humanos— en su aparato sensorial y, ciertamente, en su trasfondo cultural. El objetivo de distinguir tal representación del mundo puede ser posiblemente incoherente, pero desde luego no está motivado por la meta de trascender toda descripción y conceptualización.

No quiero seguir cuestionando hoy si es coherente la idea de un saber absoluto.[118] Menciono el asunto porque pienso que la posición de Putnam, aun en su lado equivocado, puede ayudarnos a ubicar un científicismo en la filosofía, que él y yo, en realidad, estamos de acuerdo en rechazar. El argumento básico de Putnam contra la idea del saber absoluto es que las relaciones semánticas son normativas, y, por ende, no pueden figurar en ninguna concepción puramente científica, pero describir el mundo

requiere desplegar términos que tienen relaciones semánticas con él: de ahí viene, parece, la conclusión de Putnam de que el saber absoluto supuestamente describe el mundo sin describirlo. Pasemos por alto el punto de que el argumento habla de dos cosas diferentes: por un lado, *usando* términos que tienen relaciones semánticas con el mundo, y, por el otro lado, *dando una explicación* de aquellas relaciones semánticas: me concentraré en esa segunda cuestión.[119] Si damos por aceptado, con el fin de poder argumentar, el principio, ciertamente discutible, de que las relaciones semánticas son normativas, se deduce que una explicación de ellas no puede figurar en la absoluta concepción. No se deduce que el saber absoluto sea imposible, sino solamente que una explicación de las relaciones semánticas, en particular una dada por la filosofía del lenguaje, no sería parte del saber absoluto. Pero —regresando por el momento al aspecto puramente *ad hominem* del argumento— nunca afirmé que lo sería, y en un asunto relacionado, dije que, incluso si el saber absoluto fuera alcanzable y constituyera un conocimiento del mundo “sea como sea”, sería extremadamente dudoso que pudiéramos saberlo.[120]

Entonces, ¿por qué supone Putnam, como obviamente hace, que si hubiera un saber absoluto del mundo, la filosofía tendría que ser parte de él? Dudo que lo haya desprendido simplemente de las asociaciones hegelianas de la palabra “absoluto”, con su implicación de que si hay un conocimiento absoluto, entonces la filosofía lo posee. Lo que quizá piensa es la conjunción de dos cosas: primero, que la filosofía es tan buena como puede serlo, y no es en modo alguno inferior a la ciencia, y, segundo, que si hubiera un saber absoluto del mundo, una representación del mismo que fuera lo más independiente posible de las perspectivas particulares, sería mejor que las representaciones más limitadas o condicionadas localmente. Ahora bien, la primera de estas suposiciones es, se puede decir, semiverdadera: aunque la filosofía es peor que las ciencias naturales en algunas cosas, como el descubrimiento de la naturaleza de las galaxias (o, si estaba en lo correcto acerca del saber absoluto, en la representación del mundo como es en sí mismo), es mejor que éstas en otras cosas, por ejemplo, para darle sentido a lo que tratamos de hacer en nuestras actividades intelectuales. Sin embargo, también he atribuido a Putnam la segunda suposición, de que si hubiera un saber absoluto, éste sería siempre mejor que las representaciones más particulares, y esta última es simplemente falsa. Incluso si fuera posible dar una explicación del mundo que no tuviera nada o casi nada de perspectivas particulares, no sería especialmente útil para muchos de nuestros propósitos, como entender el sentido de nuestras actividades intelectuales o de otro tipo, o incluso continuar con la mayoría de aquellas actividades. Para aquellos propósitos —en particular, al buscar entendernos a nosotros mismos—, necesitamos conceptos y explicaciones que estén enraizados en nuestras prácticas más locales, nuestra cultura y nuestra historia, y éstas no pueden ser remplazadas por conceptos que pudiéramos compartir con otros investigadores de otras partes del mundo, muy diferentes a nosotros. La ambigua palabra “nosotros” aquí no se refiere a una reunión puramente abstracta de seres cualesquiera con los cuales los seres humanos pudieran comunicarse acerca de la naturaleza del mundo, sino al conjunto de los humanos contrastados con otros seres posibles, y en el caso de muchas prácticas

humanas puede, por supuesto, referirse a agrupaciones más pequeñas.

Para resumir esta parte del argumento, hay dos errores que podemos ver. Uno es suponer que justo porque todas nuestras concepciones son nuestras, son todas igualmente locales o particulares, y que no es concebible trazar ningún contraste en este aspecto en nuestro pensamiento entre, por ejemplo, los conceptos de la física y los de la política o la ética. El otro error es suponer que si hay tal contraste, y si un conjunto de estos conceptos, por ejemplo los de la ciencia física, son potencialmente universales en su comprensión y utilidad, entonces éstos son de alguna manera intrínsecamente superiores a las concepciones más locales, que sólo están humana y quizá históricamente fundamentadas. Lo segundo es un error cientificista, y permanecerá siéndolo incluso si se niega que ese contraste sea concebible. Las personas que niegan el contraste, pero se adhieren al error —es decir, que creen que no puede haber ningún saber absoluto, pero que si lo hubiera sería mejor que cualquiera otra representación del mundo—, son a su pesar cientificistas: del mismo modo en que un ateo es realmente religioso si piensa que, puesto que Dios no existe, todo está permitido.

Debido a que Putnam supone que, si hubiera algo como la absoluta concepción del mundo, la explicación de las relaciones semánticas tendría que ser parte de ella, él también considera que es cientificista el programa filosófico, que ha adoptado varias formas, de tratar de dar una explicación de las relaciones semánticas, como la referencia, en términos científicos no normativos. Podría pensarse que es cuestionable si tal programa sería necesariamente cientificista, independientemente de las razones particulares de Putnam para pensarlo, pero de hecho me parece que esta cuestión está mal planteada. El asunto no es que el programa sea es cientificista, sino que sus motivaciones lo son, y esto mismo es una pregunta menos que clara. Doy por sentado que cualquier intento de *reducir* las relaciones semánticas a conceptos físicos está condenado al fracaso. Si, como una reacción a eso, la cuestión se reduce a que si la descripción de relaciones semánticas es consistente con la física, la contestación tiene que ser “sí”. De modo que cualquier pregunta interesante en esta área parece ser algo así como: ¿en qué medida podría identificarse el comportamiento de una criatura como comportamiento lingüístico, por ejemplo, el de referirse a algo, si dicha criatura no pertenece a un grupo que tenga algo así como como una cultura o un conjunto general de reglas que la gobiernen? Algunas preguntas se relacionan con lo anterior: ¿es el lenguaje una actividad específicamente humana, por lo que concierne a las especies terrestres, en el sentido de que está necesariamente vinculado con el rango plenamente humano de las actividades culturales autoconscientes? Nuevamente, ¿en qué etapa de evolución homínida podríamos concebir que emerge el comportamiento genuinamente lingüístico? Me parece que éstas son cuestiones realmente interesantes, y ni ellas ni su motivación son cientificistas. Lo que sería cientificista sería la suposición *a priori* de que tuvieran cierta respuesta específica, a saber, una que identificara el comportamiento lingüístico con las actividades culturales humanas en general, o que adoptara la línea reduccionista alternativa de que las actividades culturales deban ser total o mayoritariamente explicadas por la selección natural. No trataré de decir aquí nada más acerca de este aspecto del

tema, excepto para repetir una vez más la obviedad de que no son, en general, las prácticas humanas culturales las que son explicadas por la selección natural, sino más bien la característica humana universal de tener prácticas culturales, y la capacidad de los seres humanos de tenerlas. Es precisamente el hecho de que las variaciones y los desarrollos en las prácticas culturales *no* estén determinadas por la evolución lo que hace a la característica humana de vivir en culturas un éxito evolutivo tan extraordinario.

4

¿Cuáles son las tentaciones del cientificismo? Hay varias, y muchas de ellas se pueden dejar a la sociología de la vida académica, pero doy por sentado que sus motivaciones más básicas están vinculadas con la autoridad intelectual de la filosofía. La ciencia parece poseer autoridad intelectual, y la filosofía, consciente de que, como usualmente se hace, no tiene autoridad científica, puede decidir tratar de participar en ella. Ahora bien, es válido preguntar si la autoridad intelectual de la ciencia no está atada a sus esperanzas de ofrecer un saber absoluto del mundo, como es independiente de cualquier perspectiva local o peculiar sobre el mismo. Muchos científicos lo piensan. Algunas personas consideran que ésta es la única autoridad intelectual que hay. Entre éstos están incluidos, a su pesar, los defensores de las humanidades, mal dirigidos desde mi punto de vista, que piensan que tienen que mostrar que nadie tiene ninguna esperanza de ofrecer tal saber, ni siquiera los científicos: que la ciencia natural es simplemente otra parte de la conversación humana, de modo que, dejando de lado la pequeña diferencia de que las ciencias entregan refrigeradores, armas, medicinas, etc., se hallan en el mismo campo que las humanidades.[121]

Me parece que esta manera de defender las humanidades está extraviada en dos aspectos. El político, porque si la autoridad de las ciencias estuviera divorciada de cualesquiera pretensiones de ofrecer un saber absoluto, su autoridad meramente se cambiaría al hecho manifiesto de sus éxitos predictivos y tecnológicos, independientemente de la cuestión de donde provienen, y las humanidades, una vez más, en esa medida, estarían en desventaja. Intelectualmente extraviado, también por la misma razón que ya hemos encontrado, que ofrecer un saber absoluto es lo que vale la pena, lo que realmente importa para tener una autoridad intelectual, pero simplemente no hay ninguna razón para aceptar eso. Una vez más, se nos deja solos con la cuestión de entendernos mejor a nosotros mismos y a nuestras actividades, y esa cuestión incluye la pregunta, y desde luego se enfoca en ella, de cómo es que las humanidades nos ayudan a hacerlo.

Es de particular interés, desde luego, cómo hacer el mejor uso de la actividad misma de la ciencia. Aquí la historia empieza a cobrar importancia. La ciencia no nos dice gran cosa de su propia historia, y esto es una característica significativa de su práctica. (No es sorprendente que los filósofos cientificistas quieran que la filosofía la siga en esto: que piensen, como un filósofo que conozco ha dicho, que la historia de la filosofía ya no es

más parte de la filosofía que la historia de la ciencia parte de la ciencia.) Desde luego, los conceptos científicos tienen una historia: pero en la concepción estándar, aunque la historia de la física pueda ser interesante, no tiene efecto en el entendimiento de la física misma; es simplemente parte de la historia del descubrimiento.

Es, desde luego, válido preguntar qué se requiere para que una historia sea la de un descubrimiento. Una condición para eso yace en la idea familiar que formularía así: la teoría o (más generalmente) perspectiva posterior, se comprende a sí misma, a la perspectiva anterior y a la transición de ésta a aquélla, en tales términos que ambas partes (los que sostienen la perspectiva anterior, y los que sostienen la posterior) tienen razones para reconocer que la transición es una mejoría. Llamaré *vindicativa* una explicación que satisface esta condición. En el caso particular de las ciencias naturales, la teoría posterior suele explicar en sus propios términos las apariencias que apoyaron la teoría anterior, y, más aún, la teoría anterior puede ser entendida como un caso especial o limitado de la posterior, pero —y éste es un punto importante— la idea de que la explicación de la transición de una perspectiva a otra es “vindicativa”, no es definida de tal modo que se aplique solamente a las investigaciones científicas.

Los escépticos de las pretensiones de la ciencia de estar acercándose hacia un saber absoluto del mundo frecuentemente basan sus dudas en la historia de la ciencia. Niegan que la historia sea realmente vindicativa, o, en la medida en que lo es, niegan que esto sea tan importante como suele suponerse. No trataré de llevar estos argumentos más lejos, aunque es quizá digno de notarse que quienes simpatizan con este escepticismo deben ser cuidadosos con la forma en que expresan sus conclusiones históricas. Sea cual sea la concepción que se adopte de la empresa científica, uno se debería resistir a decir, como un historiador de la ciencia ha declarado sin cautela, que “la realidad de los quarks nació con la práctica de los físicos de partículas” (los años setenta son un tiempo más bien tardío para que comience el universo).[122]

5

En cualquier caso, la filosofía tiene una historia, que no es notablemente vindicativa. Me concentraré en esta parte de mi análisis en conceptos éticos y políticos, aunque muchas de las consideraciones que haré podrían ampliarse más. Si preguntamos por qué usamos algunos conceptos de este tipo más que otros —más que, digamos, los que eran usuales en un tiempo anterior—, podemos desplegar argumentos para justificarnos: ideas de igualdad y derechos iguales, por ejemplo, contra ideas jerárquicas. Por otro lado, podemos reflexionar sobre un relato histórico del modo en que estos conceptos llegaron a nosotros más que los otros: una historia (simplemente por darle una etiqueta) de cómo el mundo moderno y sus expectativas especiales llegaron a remplazar el *Ancien Régime*. Pero entonces estaremos reflexionando sobre la relación entre esta historia y los argumentos que desplegamos contra las concepciones anteriores, y nos daremos cuenta de que la historia es la historia misma de esas formas de argumento, a las que podríamos

llamar liberales y que son una parte central de nuestra perspectiva.

Si consideramos que estas ideas llegaron a prevalecer, podemos, desde luego, verlas como triunfantes, pero no necesariamente como un argumento que triunfó. Para que las ideas liberales hayan ganado una discusión, los representantes del *Ancien Régime* hubieran tenido que compartir con los nacientes liberales la idea de cuál era el punto a discutirse, y no simplemente la obviedad de que era el modo de vivir o de ordenar a la sociedad. Hubieran tenido que estar de acuerdo en que había alguna meta, de razón, libertad o lo que fuera, a la que las ideas liberales servían mejor o de las cuales eran una mejor expresión, y no hay muchas razones en un cambio tan radical, para pensar que lo estuvieran, al menos no hasta tarde en el proceso. Las ideas mismas de libertad, razón, etc., estuvieron relacionadas con el cambio. Si en este sentido los liberales no ganaron una discusión, entonces las explicaciones del modo en que el liberalismo llegó a prevalecer —es decir, entre otras cosas, cómo es que éstas vinieron a ser nuestras ideas— no son vindicativas.

El punto se puede formular también así. En el caso del cambio científico, esto puede ocurrir a través de una crisis. Si hay una crisis de explicación, las partes están de acuerdo en que la hay, al mismo tiempo que pueden, desde luego, estar en desacuerdo sobre lo que podría considerarse como una explicación. En una medida considerable tiene que llegar a haber un acuerdo, al menos dentro de los límites de la ciencia desde el siglo XVIII, y esto contribuye de modo importante a que su historia sea vindicativa. Sin embargo, las crisis que impulsaron los varios procesos geográficamente extendidos y de larga duración por los cuales el antiguo orden político y ético se convirtió en modernidad, no fueron en primera instancia crisis de explicación; fueron crisis de confianza o de legitimidad, y la historia del modo en que una concepción, más que otra, llegó a proveer la base de una nueva legitimidad, no es a primera vista vindicativa.

Hay, desde luego, o ha habido, historias que tratan de vindicar históricamente una u otra concepción moderna, como un desarrollo de la razón o de la ilustración, o como una realización más plena de la libertad y la autonomía, que es un objetivo humano constante, y hay otras más. Tales historias son poco populares por el momento, particularmente en las versiones de pantalla ancha ofrecidas por Hegel y Marx. Entre los filósofos de nuestra tradición local, esas historias son poco populares no tanto en el sentido de que sean negadas, sino de que ni siquiera se mencionan. Sin duda, no se mencionan, en parte, porque no se cree en ellas, pero también porque no es parte de una empresa filosófica, como se entiende localmente, poner atención a cualquier historia como ésa, pero —y éste es el punto que quiero enfatizar— *debemos* ponerles atención, si es que hemos de saber qué actitud reflexiva adoptar hacia nuestras propias concepciones. Una razón de eso es que responder si hay o no una historia de nuestras concepciones que sea vindicativa (aunque sea modestamente), es importante para poder decir, si es que lo decimos, que las concepciones anteriores estaban equivocadas. En ausencia de explicaciones vindicativas, mientras que es posible, por supuesto, decir que las viejas ideas estaban equivocadas —¿quién puede impedirlo?—, el contenido de esto es probablemente muy pobre: transmite solamente el mensaje de que la perspectiva anterior

fracasó frente a argumentos basados en que tales perspectivas fracasaran por ellos. Es una buena pregunta si una melodía tan pobre como ésta es digna de ser siquiera silbada.

Sin embargo, este asunto (la cuestión, a grandes rasgos, del relativismo) no es lo principal. Lo que en verdad nos ocupa aquí es nuestra actitud filosófica hacia *nuestras propias* concepciones. Incluso fuera de la vindicación y las consecuencias que ésta pueda tener al comparar nuestras perspectivas con otras, los filósofos no pueden ignorar enteramente la historia si pretender entender nuestros conceptos éticos. Una razón de eso es que en muchos casos el contenido de nuestros conceptos es un fenómeno histórico contingente, lo que se debe a más de una razón. Por tomar un caso en el que estoy actualmente trabajando, las virtudes asociadas con la veracidad, pienso que está claro que, mientras que hay una necesidad humana universal de cualidades como la precisión (la disposición a adquirir creencias verdaderas) y la sinceridad (la disposición a decir lo que se cree que es verdadero), las formas de estas disposiciones y de las motivaciones que encarnan son cultural e históricamente diferentes. Si se quiere entender nuestra propia concepción de tales cosas, y hacerlo en términos que puedan considerarse filosóficos —por ejemplo, para salir de dudas acerca de la base de estos valores y sus implicaciones—, debe tratar de entenderse por qué entre nosotros adoptan ciertas formas más que otras, y sólo puede hacer eso con la ayuda de la historia. Más aún, hay algunas virtudes como la autenticidad o cierto tipo de integridad, que son en su totalidad un desarrollo cultural manifiestamente contingente; no hubieran evolucionado en absoluto si la historia de occidente no hubiera seguido cierto curso. Por estas dos razones, el entendimiento reflexivo de nuestras ideas y motivaciones, que supongo que es por acuerdo general una meta filosófica, requiere entendimiento histórico. Aquí la historia ayuda al entendimiento filosófico, o es parte de él. La filosofía tiene que aprender que la descripción conceptual (o, más específicamente, el análisis) no es autosuficiente, y que el intento de derivar nuestros conceptos *a priori* de condiciones universales de la vida humana, aunque, desde luego, tiene un lugar (un lugar más grande en algunas áreas de la filosofía que en otras), probablemente deje sin explicar muchas motivaciones de la investigación filosófica.

6

Hay otros asuntos, sin embargo, en los que puede parecer que el entendimiento histórico no ayuda a la empresa filosófica, sino que más bien la estorba. Si pensamos que nuestra perspectiva tiene una historia vindicativa, podríamos en esa medida ignorarla, precisamente como los científicos ignoran la historia de la ciencia. (Pueden verse las enormes e implausibles suposiciones de quienes piensan que la filosofía puede ignorar su propia historia.) Pero si no creemos que la historia de nuestra perspectiva es vindicativa, puede parecer que entenderla interfiere con nuestro compromiso con dicha perspectiva, y en particular con un intento filosófico de trabajar dentro de ella y desarrollar sus argumentos. Si lo que tiene lugar aquí y ahora es un desarrollo contingente, ¿podemos

nosotros identificarnos plenamente con él? ¿Es realmente *nuestro*, excepto en el sentido de que él y nosotros estamos en el mismo lugar al mismo tiempo?

En alguna medida, ésta es una versión de un problema recurrente en el pensamiento europeo desde que la autoconciencia histórica echó raíces profundas a principios del siglo XIX: un problema de reflexión y compromiso, o de una concepción externa de las creencias propias como opuesta a una profundización interna en ellas; un problema, como podría llamarse, de fatiga historicista y alienación. Puede ser un testimonio del poder de este problema que tantos filósofos liberales quieran evitar cualquier cuestión relativa a la historia de sus propias concepciones. Puede ser también importante en esta conexión que tanta filosofía política robusta e influyente provenga de los Estados Unidos, que no emergió históricamente del *Ancien Régime*, aunque (por decirlo así) procede de él por el mero acto de aparecer en la historia.

Un filósofo, desde luego estadounidense, que ha planteado la pregunta dentro de la tradición local, es Richard Rorty, quien ha sugerido que la respuesta a ello se halla en la ironía:[123] que *qua* actores políticos estamos enfocados en cierta perspectiva, pero *qua* personas reflexivas (por ejemplo, como filósofos) damos un paso atrás y, con un espíritu desapegado y más bien irónico nos vemos a nosotros mismos con ese desapego. El hecho de que ese *qua* deba venir tan naturalmente al formular esta perspectiva, muestra, como casi siempre en filosofía, que se está tratando de separar lo inseparable: en este caso, lo éticamente inseparable, y probablemente también lo psicológicamente inseparable, a menos que el ironista se una a los otros (la perspectiva que Rorty llama “sentido común”) y se olvide enteramente del autoentendimiento histórico, en cuyo caso puede también olvidar la ironía, y, desde luego, no la necesita.

De hecho, me parece que una vez que se va *lo suficientemente lejos* al reconocer la contingencia, el problema que la ironía supuestamente iba a solucionar ya no surge en absoluto. Lo que tenemos aquí es muy similar a algo que ya hemos encontrado en este análisis: el fenómeno del cientificismo involuntario. El supuesto problema proviene de la idea de que queramos tener una historia vindicativa de nuestra perspectiva, y es una decepción el descubrimiento de que el liberalismo en particular (aunque lo mismo es cierto de cualquier perspectiva), tiene el tipo de historia contingente que tiene, lo que nos deja en el mejor de los casos un segundo mejor. Pero una vez más, ¿por qué deberíamos pensar eso? Precisamente porque no somos de inteligencias plenamente libres, seleccionando en principio entre todas las perspectivas posibles, podemos aceptar que esta perspectiva es nuestra justamente porque la historia la ha hecho nuestra, o más precisamente, en la medida en que la ha hecho nuestra, así como ha hecho la perspectiva como algo que es nuestro. No estamos menos contingentemente formados que la perspectiva en cuestión, y la formación es significativamente la misma. Nosotros y nuestra perspectiva no estamos simplemente en el mismo lugar y el mismo tiempo. Si realmente entendemos esto, si lo entendemos profundamente, podemos ser libres de lo que es, desde luego, otra ilusión cientificista, la idea de que es nuestra tarea como agentes racionales buscar, o al menos acercarnos en la medida de lo posible, hacia un sistema de ideas políticas y éticas que sería mejor desde un punto de vista absoluto, un punto de

vista que esté libre de perspectivas históricas contingentes.

Si podemos deshacernos de esa ilusión, veremos que no hay un conflicto inherente entre tres actividades: primero, la actividad de primer orden de actuar y argumentar dentro del marco de nuestras ideas; segundo, la actividad filosófica de reflexionar sobre aquellas ideas en un nivel más general y tratar de entenderlas de la mejor manera, y tercero, la actividad histórica de entender de dónde provienen. Esas actividades tienen de varias maneras continuidad entre sí. Esto ayuda tanto a conseguir inteligencia en la acción política (debido a la conexión de la primera con la segunda y la tercera), como realismo en filosofía política (debido a la conexión de la segunda con la primera y la tercera). Si hay una dificultad al combinar la tercera de estas actividades con las primeras dos, es pensar sobre dos cosas al mismo tiempo, no tomar en serio a ambas a la vez.

7

De hecho, es muy improbable que seamos capaces de comprender completamente nuestra propia perspectiva. Sería incoherente por varias razones. La historia puede ayudarnos a entender por qué eso debe ser así; por ejemplo: las dificultades que el liberalismo tiene actualmente con las ideas de autonomía se pueden retrotraer en parte a las concepciones del individuo de la Ilustración, que ahora no tienen sentido pleno para nosotros. En estas circunstancias podemos, desde luego, estar separados de parte de nuestra propia perspectiva. Si la incoherencia es lo suficientemente severa, se nos presentará a quienes mantenemos esta perspectiva como una crisis de explicación: necesitamos tener razones para reordenar y desarrollar nuestras ideas de un modo más que de otro. Al mismo tiempo, podemos quizá ver la situación como una crisis de legitimidad: que se dude seriamente de si estas ideas sobrevivirán y continuarán sirviéndonos. Otros que no comparten la perspectiva en cuestión, pueden ver también la crisis de legitimidad, pero no pueden verla como una crisis de explicación para sí mismos, puesto que ellos, en primer lugar, no pensaban que nuestra perspectiva explicara las cosas. Nosotros, sin embargo, necesitamos razones internas a nuestra perspectiva no simplemente para encontrar explicaciones, sino también para evitar una crisis de legitimidad. También las necesitamos para explicarnos a nosotros mismos frente a personas que están indecisas entre nuestra perspectiva presente y la de algún rival activo contemporáneo. Si las cosas están lo suficientemente mal, nosotros mismos podemos formar parte de esas personas.

Puede no haber crisis, o si la hay, algunos elementos de nuestra perspectiva permanecerán juntos. Creemos, por ejemplo, que en algún sentido todo ciudadano, todo ser humano —algunas personas, de un modo más extravagante, dirían que todo ser sensible—, merece igual consideración. Quizá esto es menos una creencia proposicional que el esquema de varios argumentos, pero en cualquier caso puede parecer, al menos en su forma más central e inespecífica, *unhintergebar*: no hay forma más básica de justificarla. Sabemos que la mayoría de las personas en el pasado no la han compartido;

sabemos que hay otros en el mundo que no la comparten ahora, pero para nosotros está simplemente ahí. Eso no significa que pensemos: “para nosotros, está simplemente ahí”. Significa que pensemos: “está simplemente ahí”. (Eso es lo que significa que esté, para nosotros, simplemente ahí.)

Al menos con respecto a estos elementos de nuestra perspectiva, un filósofo puede decir: la historia contingente no tiene efectos en el espacio de las razones (para usar una frase de moda), entonces ¿por qué preocuparse por ella?[124] Continuemos con nuestro trabajo de explicar nuestra perspectiva desde adentro lo mejor que podamos. Hay varias respuestas a esto, algunas implícitas en lo que ya he dicho. Una es que los filósofos reflexionando sobre estas creencias o modos de argumentación pueden retornar a esos antiguos artilugios de reafirmación cognitiva como la “intuición”, pero si las aseveraciones epistémicas implícitas en tales términos han de ser tomadas seriamente, entonces tienen implicaciones en la historia —implican una historia *diferente*—. Nuevamente, lo que pensamos acerca de estas cosas afecta nuestra forma de ver a las personas que en el presente tienen diferentes perspectivas, que se presentan como rivales a las nuestras. Decir simplemente que en nuestros términos estas personas están equivocadas es retornar a la pobre tonada que ya hemos escuchado en el caso de la desaprobación a lo largo de los siglos. Es importante por qué estas personas creen lo que creen; por ejemplo, si podemos razonablemente considerar su perspectiva simplemente como arcaica, como una expresión de un orden que ha sobrevivido en un ambiente internacional en el que no puede durar, social o intelectualmente, entonces eso influye no sólo en la persuasión de las partes no comprometidas, como ya he dicho, sino también en la explicación de la relación de otros con nosotros mismos, y, por ende, de nosotros mismos en relación con ellos. Incluso, con respecto a los elementos de nuestra perspectiva para los cuales no hay ulteriores justificaciones, puede aún haber explicaciones que ayuden a ubicarlas en relación con sus rivales.

Por encima de todo, el entendimiento histórico —quizá pueda decir ahora, más ampliamente, el entendimiento social— puede ayudar en la labor, que es ciertamente filosófica, de distinguir entre los diferentes modos en que varias de nuestras ideas y procedimientos pueden parecer insuperables, sin una alternativa concebible.

Esto nos regresa a Wittgenstein, que insistió correctamente y de modo influyente en que las justificaciones tenían un final, que en varios puntos encontramos el hecho de que “éste es el modo en que continuamos”. Pero, si puedo decir nuevamente algo que ya he dicho antes, más bien frecuentemente,[125] hace una gran diferencia quién se supone que somos “nosotros”; con esa palabra podemos referirnos a grupos diferentes en diferentes conexiones filosóficas. Puede referirse, al extremo, como mencioné anteriormente, a cualquier criatura imaginable dotada de entendimiento. O puede referirse a cualquier ser humano, y aquí pueden ser relevantes las condiciones universales de la vida humana, incluidas capacidades psicológicas muy generales. O puede simplemente referirse a aquellos con quienes usted y yo compartimos más, como las perspectivas típicas de la modernidad. El mismo Wittgenstein heredó de Kant una preocupación por los límites del entendimiento; de Frege y Russell un interés en las condiciones del significado lingüístico,

y de sí mismo un sentido de la filosofía como una empresa muy peculiar y posiblemente patológica. Estas influencias lo guiaron hacia las cuestiones más generales de la filosofía, y, con eso, a un entendimiento más amplio de “nosotros”, pero también conspiraron para hacerle pensar que la filosofía no tenía nada que ver con las explicaciones, no sólo las explicaciones científicas (ciertamente fue el menos científicista de los filósofos), sino con cualquier clase de explicaciones, salvo las filosóficas, y estas últimas no eran para él como las demás explicaciones, sino más bien como elucidaciones o recordatorios. En este sentido, sus modos de hacer filosofía, y, desde luego, sus dudas acerca de ella, todavía se enfocaban en una concepción de la filosofía como exclusivamente *a priori*. Tenemos buenas razones para cuestionar esta concepción, y él, desde luego, así lo hizo.

Una vez que abandonamos esa suposición, podemos adoptar un interés filosófico legítimo en lo que sea aceptado como un “nosotros” más local, pero se puede decir que, cuando es específicamente este grupo más restringido el que está en cuestión, no puede ser que no tenga alternativas concebibles. ¿Es seguro que la historia con la que he estado tratando es una historia de alternativas? Pero eso es un mal entendido de lo que, en este contexto, se dice que es inconcebible. La historia presenta alternativas solamente para un “nosotros” más amplio: presenta modos diferentes, en que los seres humanos han vivido y, por ende, pueden vivir. Claramente, en aquellos términos podemos ser capaces de concebir, si bien sólo esquemáticamente y con dificultad, otros modos en que los seres humanos podrían vivir en el futuro, pero ese no es el punto. Lo que en esta conexión “parece simplemente estar ahí”, y no tener alternativa, son ciertos elementos de nuestra perspectiva ética y política, y en ese sentido no habría alternativas para nosotros. Esos términos son, desde luego, *unhintergebar*, en un sentido que, por supuesto, implica tiempo, pero en una forma peculiar. Podemos explorarlos por este lado, en relación con su pasado, y explicarlos, y (si, como ya he dicho, abandonamos las ilusiones científicistas) podemos identificarnos con el proceso que condujo a nuestra perspectiva porque podemos identificarnos con su resultado, pero no podemos en nuestro pensamiento ir más allá de nuestra perspectiva hacia el futuro y permanecer identificados con el resultado: es decir, no podemos superar nuestra perspectiva. Si un futuro posible que figura en aquellas vagas especulaciones no encarna alguna interpretación de estos elementos centrales de nuestra perspectiva, entonces puede tener sentido empírico para nosotros —podemos ver cómo alguien pudiera llegar allí—, pero no ético, excepto como una ciencia de regresión, desolación, pérdida.

Eso está conectado con el hecho de que las concepciones modernas éticas y políticas no suelen aceptar un futuro que las rebase. El marxismo predijo un futuro que, se suponía, tenía sentido ético, pero notoriamente llegó a un fin en una utopía estática. Muchos liberales a su manera siguieron el mismo patrón; procedieron en esta cuestión como el pasado, como si el liberalismo fuese atemporal.[126] No puede reprocharse a estos liberales que no puedan ver más allá de los límites externos de lo que encuentran aceptable: nadie puede hacer eso, pero es más que reprochable que no estén interesados lo suficientemente en la razón de que esto sea así, en la razón de que deba parecer que sus convicciones más básicas estén, como lo dije, “simplemente allí”. Es parte de una

actitud filosófica la que los hace estar igualmente desinteresados en el modo en que estas convicciones llegaron allí.

8

He argumentado que la filosofía debería deshacerse de ilusiones científicistas, que no debería tratar de comportarse como una extensión de las ciencias naturales (excepto en los casos especiales donde eso es lo que es), que debería pensar en sí misma como parte de una empresa humanística más amplia de explicarnos a nosotros mismos y nuestras actividades, y que para responder a muchas de sus preguntas necesita poner atención a otras partes de esa empresa, en particular a la historia.

Pero alguien, quizá un joven filósofo, podría decir: todo eso está muy bien, pero incluso si lo acepto completamente, no significa que haya demasiado que necesito saber, ¿que uno solamente puede hacer filosofía siendo absolutamente un amateur? ¿No podemos simplemente proceder con ello?

A él solamente le puedo decir: entiendo perfectamente tu, es decir nuestro, problema. Acepto que la filosofía analítica debe mucho de su éxito al principio de que poco y bueno es mejor que mucho y malo; acepto que esto implica una división del trabajo; acepto que tú quieres continuar con ella. También admito algo más, que son comúnmente los filósofos maduros quienes, como los científicos maduros, tienden a reflexionar en estos modos tan amplios acerca de la naturaleza de su materia. Como dice Nietzsche en un pasaje maravilloso acerca del filósofo y la edad:[127]

Muy frecuentemente sucede que el viejo está sujeto a la falsa ilusión de una gran renovación y renacimiento moral, y a partir de esa experiencia emite juicios sobre la obra y el curso de su vida, como si apenas ahora hubiera adquirido una visión clara; y, no obstante, la fuente de inspiración que está detrás este sentimiento de bienestar y estos juicios confiados no es la sabiduría, sino la *fatiga*.

Sin embargo, se deben decir cosas acerca de las formas en que uno podría aceptar la perspectiva de la filosofía que estoy ofreciendo, y, no obstante, continuar practicándola. Déjenme terminar mencionando muy brevemente una o dos de ellas. Una cosa que necesitamos hacer no es abandonar la división del trabajo, sino reconsiderarla. Tiende a ser moldeada demasiado fácilmente sobre la de las ciencias, dividiendo un campo o área del teorizar de otra, pero podemos dividir la materia de otros modos: pensando en una idea ética dada, por ejemplo, y en las varias consideraciones que pudieran ayudarle a uno a entenderla. Nuevamente, mientras que es ciertamente verdadero que todos necesitamos saber más de lo que podemos esperar saber —y eso es verdad de los filósofos que trabajan cerca de las ciencias o, desde luego, de los que trabajan en ellas— hace una gran diferencia qué es lo que se sabe que no se sabe. Uno puede no ver muy lejos fuera de la propia casa, pero puede ser muy importante en qué dirección está uno buscando.

Más aún, no es meramente un asunto de investigación o de escritura filosófica, sino también de la impresión que se da de la materia al enseñarla. La mayoría de los

estudiantes no tienen interés de convertirse en filósofos profesionales. Frecuentemente, se llevan una imagen de la filosofía como una materia técnica autocontenida, y se admite que esto puede tener su propio encanto como algo complicado que puede ser hecho bien o mal, lo que no ha de ser despreciado. Esto también de cierto modo, hace que la materia sea más fácil de enseñar, puesto que requiere menos tratar de encontrar qué tanto saben los estudiantes acerca de cualquier otra cosa. Pero si creemos que la filosofía podría contribuir de manera importante a hacer que las personas piensen lo que están haciendo, entonces dicha disciplina debería reconocer sus conexiones con otros modos de entendernos a nosotros mismos, y si continúa resistiéndose a hacerlo, puede parecer al estudiante algo muy extravagante desde cualquier punto de vista.

Corremos el riesgo, de hecho, de que la entera empresa humanística de tratar de entendernos esté llegando a parecer extraña, y eso por varias razones: la educación está concentrándose en lo técnico y lo comercial, hasta un punto en el que cualquier investigación más reflexiva puede parecer innecesaria y arcaica, algo que en el mejor de los casos es preservado como parte de nuestra herencia. Si es así como se preserva, no será la actividad apasionada e inteligente que necesita ser. Todos tenemos un interés en la vida de esa actividad —no simplemente un interés compartido, sino un interés en una actividad compartida.

XVII. ¿En qué podría convertirse la filosofía?

1

Si hemos de tener etiquetas en filosofía, éste es un buen momento para crear una nueva, y “filosofía postanalítica” es una etiqueta atractiva.[128] Sin embargo, inmediatamente plantea una o dos cuestiones de política intelectual.

Primero, estoy seguro de que no debería perderse tiempo tratando de definir lo “analítico” en cuanto a que la filosofía postanalítica sea “post”. Ello podría regresarnos a controversias estériles reminiscentes de los volúmenes suplementarios de *Proceedings of the Aristotelian Society* de la década de 1930, con títulos como “¿Es la filosofía el análisis del sentido común?”, y esto sería de lo más inapropiado cuando la distinción entre analítico y sintético, sobre lo cual presumiblemente descansaba la noción de análisis, es peor considerada. Más aún, concentrarse en esa particular noción ocultaría poderosos recursos que han sido desplegados por la misma filosofía analítica, en cualquiera de los sentidos que ordinariamente se han entendido. Un ejemplo que ilustra explícitamente el punto, es el libro de Edward Craig, *Knowledge and the State of Nature*, [129] que está sarcásticamente subtítuloado *Un ensayo en síntesis conceptual*. El método de Craig es desplegar una historia de desarrollo ficticia (más aún, imposible), más bien un conjunto de historias desplegadas en la filosofía política tradicional acerca del “origen” del Estado. Las genealogías ficticias, como yo las llamaría, pueden ser muy ilustrativas, y, como en el caso de Craig, pueden pertenecer a lo que es inequívocamente filosofía analítica, al mismo tiempo que ofrecen deliberadamente un intento de dar una alternativa al análisis tradicional. No trataré de decir aquí nada más acerca de las genealogías ficticias y lo que pueden ofrecer, aunque regresaré posteriormente al papel que pueden tener las genealogías reales, la historia genuina.

Segundo, no queremos que “postanalítico” se convierta en el nombre de un estilo opuesto a otros. La categoría de “nuestro estilo”, en contraposición a “los otros estilos” no tiene ninguna utilidad; sin embargo, no importa de qué manera sea exhibida, y lo es igualmente, ya sea que simplemente defendamos nuestro propio estilo y ataquemos el de otros, o tratemos de elaborar términos de coexistencia irénica [promotora de la paz] con ellos sobre una base más o menos diplomática. Si realmente hay para nosotros una cuestión acerca de ellos —con excepción de las circunstancias en las que literal y desafortunadamente se ha convertido en un asunto de política académica—, ésta sería entonces una cuestión planteada para nosotros *por* ellos; una cuestión que debiéramos

ser capaces de reconocer, de un modo o de otro, sobre el modo en que deberíamos entender lo que estamos haciendo.

Para terminar estas observaciones preliminares, quiero decir que es particularmente importante que “postanalítico” no debe ser entendido en términos de la supuesta distinción entre filosofía analítica y continental.[130] Digo esto como alguien que es, tanto como negable e innegablemente, un filósofo analítico: negablemente, porque estoy dispuesto a negarlo, e innegablemente, porque sospecho que pocos que tengan algo que decir sobre el tema aceptarán esa negación. Lo que quiero negar es la utilidad de la distinción misma, lo que marcaré, en particular, diciendo muy poco acerca de ella. Pero vale la pena enfatizar que lo que es inútil en esta distinción no es solamente la cuestión de las infortunadas etiquetas que usa.

Dichas etiquetas son doblemente desafortunadas. Primeramente, implican una clasificación cruzada entre lo metodológico y lo geográfico: es como distinguir entre autos japoneses y de tracción delantera. Pero, por añadidura, como nos lo ha recordado Michael Dummett, las consideraciones geográficas son de cualquier manera erróneas, puesto que algunas de las más importantes influencias originales sobre la filosofía analítica provienen del mundo de habla germánica.[131] Pero no debemos suponer que la distinción estaría bien si fuera simplemente reetiquetada, pues no está claro qué es lo que se supone que está en cuestión entre estos diferentes modos de proceder en (algo que pudiera ser llamado) filosofía: no es obvio qué es lo que se estaría reetiquetando, ni hasta qué punto.

Con algunos escritores contemporáneos, la cuestión no es tanto si deberían leerse, sino dónde —si en un departamento de filosofía o de literatura (aunque, desde luego, un grupo u otro podría pensar que algunos escritores no son dignos de ser leídos en absoluto)—. Hay, por supuesto, casos de escritores que claramente aparecerán en una lista y no en la otra, pero tales listas no son autoexplicativas. Más aún, pueden ser distractoras e inútiles, por las asociaciones que imponen (por ejemplo, para alguien a quien le gustaría apropiarse mucho de Nietzsche, pero no en modos que tengan mucho que ver con Heidegger). Tratar de pensar en algún principio o explicación general supuestamente implícita en estas listas, plantea una cuestión sobre la identidad profesional de la filosofía.

2

Si los miembros de un departamento de filosofía son desafiados respecto de este punto, o si se desafían a sí mismos, encontrarán y probablemente consultarán la autoimagen profesional de la filosofía. Supóngase que nuestra imagen de la filosofía es alentada por la filosofía analítica. Reflexionemos sobre las virtudes —las indudables virtudes— de tal filosofía: claridad, precisión, control (o algo como eso), posibilidades de discusión cuidadosa y cooperación entre diferentes personas. (Parte de las bromas acerca de *Limited inc*,[132] una parte que no era la intención de Derrida, fue su incompreensión de

la idea misma de que debería hacerse tal cosa). Hay un concepto de integridad profesional, que incluye todo esto, pero también la idea de que lo que la filosofía hace en este espíritu lo hace especialmente (lo hace, uno podría igualmente decir, como una especialización).

Ahora bien, ¿por qué son éstas las virtudes, o al menos las virtudes básicas y centrales, de la filosofía? Una cosa que se podría decir es que en la filosofía tiene que haber algo que pueda considerarse como una forma correcta de proceder. Eso es contundente, y de una importancia central. Tengo cierta simpatía por los intentos de Richard Rorty de cuestionar las autoimágenes de la filosofía, y de ampliar su rango de referencia, pero lo arruina todo cuando termina concluyendo, como me parece que hace, que nada puede tomarse como una forma correcta de proceder.

Hace una inferencia equivocada: de la aseveración de que no hay una definición interesante de verdad (lo cual es verdadero), a través de la aseveración de que el concepto de verdad no es tampoco muy interesante (lo cual es al menos discutible), a la conclusión de que tratar de decir algo verdadero no es tan interesante como tratar de decir algo interesante, lo cual es falso: en particular, porque acerca de las cosas interesantes, lo que es interesante es que son verdaderas.

Rorty favorece un modelo de *conversación*[133] para el futuro de lo que solía ser llamado filosofía; lo propone igualmente como un modelo para el futuro de otras actividades que solían ser llamadas de modos distintos, como la crítica literaria (de hecho, parece ser el modelo para casi todo, incluyendo las ciencias naturales). Pero el modelo no es alentador; a menos que una conversación sea muy implacable —por ejemplo, entre filósofos— no será cuestionada por “así”, “por lo tanto” o “pero”, sino más bien por “bien entonces”, “eso me recuerda” y “llegué a pensar acerca de ello”, y simplemente no está claro quién permanecerá alrededor de conversaciones que supuestamente heredarían el papel de la filosofía, o por qué. De hecho, es tentador pensar que el modelo de la conversación es secretamente un aliado de la profesionalización: las únicas personas que habrán de tomar parte en tal conversación son las que pagan para hacerlo. En relación con lo anterior, el modelo difícilmente puede proveer alguna base para defender la filosofía, o, más generalmente, las humanidades, de quienes quieren excluirlas de *su* conversación.

A la forma correcta de proceder se le tiene que dar su lugar, y lo mismo vale, indiscutiblemente, para la claridad y la precisión, pero hay más que un modo de conseguir todas estas cosas —por ejemplo, más de un modo de proceder que pueda considerarse correcto: depende de qué *es*—. Entonces, ¿qué es lo que significa, en relación con esta particular interpretación de “proceder correctamente”, que los escritos en el estilo analítico terminen en el plan de estudios de filosofía, mientras que los escritos realizados en algunos otros estilos filosóficos terminen en el departamento de literatura?

No estoy negando que pueda haber una respuesta a esa pregunta. Parte de lo que quiero decir es que puede haber más de una respuesta a ella, y en algunos casos dichas respuestas pueden estar justificadas. Pero lo que quiero cuestionar es la idea de que haya un estilo que defina muy clara y uniformemente a lo largo y ancho del espectro de la

filosofía qué puede considerarse como claridad, como precisión y como proceder correctamente, y que este estilo haya sido definido por los procedimientos típicos de la filosofía analítica. Es común que se tenga esta idea, pero ¿cuál es el contenido de la noción de proceder correctamente que va con esta idea, y de dónde proviene? Es difícil negar, en una parte demasiado grande del tema, que la idea de proceder correctamente que ha ido en la autoimagen de la filosofía analítica, y que ha apoyado algunas de sus exclusiones, es extraída de las ciencias naturales, y los efectos de esto pueden ser desafortunados.

3

Hay áreas de la filosofía muy cercanas a las ciencias: la filosofía de la mecánica cuántica, la lingüística, la filosofía de las matemáticas, la probabilidad, alguna filosofía de la psicología (aunque probablemente menos que lo que algunos trabajadores del campo suponen). En cualquier ciencia hay una continuidad entre problemas conceptuales, por un lado, y trabajo empírico y construcción teórica, por el otro. En estos campos particulares hay, más aún, continuidades entre los aspectos más abstractos y filosóficos de los problemas conceptuales, y los aspectos que se hallan más cercanos al trabajo empírico y la teoría científica. Así que aquí hay espacio para un teorizar filosófico responsable: algo más cercano a la teoría científica, una propuesta de reorganización de los conceptos que podría ser seriamente discutida con alguien dedicado al trabajo científico.

La filosofía ha dado a luz a varias ciencias y puede hacerlo nuevamente, pero en mucho de la filosofía estas relaciones no son posibles, e incluso en áreas de lo más relevantes, la ciencia puede ser un mal modelo para la filosofía. Podemos dejar de lado el fenómeno de la filosofía que pretende ser ciencia cuando no lo es —las afectaciones de “la filosofía del *e-mail*”, como podría llamarse, la familiar autoconciencia y febril actividad profesional que juega un juego de hallarse al filo del progreso científico empírico—. Más seriamente, hay un científicismo de estilo y procedimiento que está mucho más difundido y no tiene obstrucciones.

Hay varias características de la ciencia natural que hacen que su modelo, cuando se aplica a la filosofía, pueda tener un efecto pernicioso sobre ella. Una de estas características es que la ciencia realmente no necesita conocer su propia historia. Es sin duda deseable que los científicos supieran algo de la historia de su ciencia, pero no es esencial para sus investigaciones. Algunos filósofos han extraído una conclusión paralela: en un prestigioso departamento estadounidense, una figura de la más alta categoría tenía un aviso sobre su puerta que decía: SIMPLEMENTE DI NO A LA HISTORIA DE LA FILOSOFÍA. En una o dos áreas la filosofía puede encontrarse lo suficientemente cerca de la ciencia como para que esta actitud esté justificada, pero si ello es así, son una excepción. En general, uno debe tomar extremadamente en serio la advertencia de Santayana de que quienes ignoran la historia de la filosofía están condenados a revivirla

(no simplemente reinventar la rueda, sino a reinventar la rueda cuadrada).

Una segunda característica es que la ciencia realmente tiene una división efectiva del trabajo. Es, desde luego, cierto que se han logrado grandes progresos mediante la transferencia de habilidades entre los campos científicos: por ejemplo, por John Maynard Smith, educado como ingeniero, al cambiarse al campo de la biología. Pero en la práctica cotidiana hay métodos perfectamente bien establecidos para obtener resultados locales, e incluso si éstos no son muy emocionantes, son resultados. Se deduce de esto que los ejercicios de formación profesional, no importa cuán ordinarios sean, hacen una contribución no solamente a la educación de las personas en la materia, sino a la materia misma.

Esto no es necesariamente verdadero en la filosofía. Sería un mal amigo de ella quien no admitiera que hay mucho trabajo filosófico que no es gratificante bajo ningún estándar: inútil, aburrido, estéril. El hecho horrible es que algo de él intenta con todas sus fuerzas ser otra cosa. Consiste en ejercicios que son necesarios para la estructura de la filosofía en tanto carrera académica. La profesionalización de la filosofía ha tenido lugar por más de un siglo (o más aún, si se cuenta la edad media), pero su nivel actual no tiene precedentes. Esto innegablemente trae sus propias deformaciones, y la pregunta que Stravinski solía hacer ofensivamente acerca de mucho de la música contemporánea —“¿quién la necesita?”—, puede ser planteada en contra de muchos productos de la filosofía académica.

Finalmente, hay una cuestión de estilo escrito. La ciencia, por supuesto, exhibe imaginación,^[134] pero cuando lo hace, tiende a ser creativa más que expresiva. Conduce a un descubrimiento o una teoría, y no necesariamente emerge del modo en que está expresado el resultado. El escribir científico debería ser claro y efectivo, y puede tener estilo, pero el éxito de los científicos no es afectado por los poderes expresivos de su manera de escribir. Esto no es necesariamente así con la filosofía. Las tradiciones del estilo llano, que son familiares a la filosofía analítica, tienen mucho que decir a su favor, pero se pueden convertir en un peso muerto bajo la influencia del modelo científico. Uno no debería aproximarse a la escritura filosófica en el espíritu del filósofo analítico que (en realidad) le dijo a otro, cuando estaban tratando de escribir un libro juntos. “Primero vamos a hacerlo bien, después puedes darle estilo”.

¿Por qué deberíamos suponer que esto debe ser así? Cuando nos dirigimos, en particular, a la filosofía moral y política, y miramos el canon de la filosofía del pasado en la que incluso la filosofía analítica está de acuerdo, ¿tiene esta apariencia? Platón, Hobbes, Hume, Rousseau, desde luego John Stuart Mill, para no ir a un territorio más disputado: ¿suponemos realmente que sus contribuciones al tema son independientes de los poderes imaginativos y expresivos de su trabajo? Se halla, desde luego, el extraordinario y sin paralelo caso de Aristóteles, quien ha tenido una inmensa influencia sobre la concepción de la tradición analítica, lo que es proceder correctamente, pero ¿por qué deberíamos siquiera suponer que estos tratados carentes de afecto representan su propia voz? En la medida en que lo hacen, ¿qué significa su tono? Las imágenes que Aristóteles da o deja implícitas de la sociedad en que vivió son ficticias en un grado

notable: ¿quizá deberíamos ver en sí misma la carencia de color, de historia, el detalle técnico en sí mismos como una evasión? En cualquier caso, ¿por qué debiéramos querer sonar así? A la mayoría de los filósofos no les hace justicia su legado histórico: Platón no es verdaderamente correspondido por las corrientes platónicas, e incluso pasó lo mismo con Hegel, pero Aristóteles, quizá de manera exclusiva, merecía lo que obtuvo: inventó la escolástica.

Como aquellos otros autores (y muchos más) nos recuerdan, la filosofía moral y política (al menos) exige algo más que tal estilo. Un filósofo puede necesitar darnos una imagen de la vida, la sociedad y el individuo, y dárnosla de un modo que la integre con lo que a él le importa. Si un escritor filosófico no resuelve los problemas del modo de expresar esas preocupaciones adecuadamente, o, como en muchos casos, ni siquiera encara aquellos problemas, no habrá logrado llevar la reflexión lo suficientemente lejos. Así que la exigencia de que la filosofía moral y política suene bien, que hable con una voz real, no es algo arbitrariamente impuesto por quienes tienen un gusto particular por la literatura, la historia o la emoción. Se deduce del ideal de la reflexividad de la filosofía, un ideal reconocido por sus tradiciones más centrales.

Desde luego, hay otra cuestión acerca de la atractiva cualidad del modelo científico, donde *es* más apropiado, y acerca de la atractiva cualidad de la ciencia misma. Las versiones que algunos filósofos dan de la ciencia no explican adecuadamente su poder moral como paradigma del ideal de veracidad; un ideal que frecuente y apasionadamente es evocado al reivindicar la aplicación de ese paradigma al trabajo filosófico en el estilo analítico. Como lo dijera Nietzsche, el motivo subyacente de esto no es “no habré de ser engañado”, sino “no habré de engañar, ni siquiera a mí mismo”, “y con ello estamos sobre un terreno moral”. [135]

Las metas de la filosofía, he sugerido, requieren que el trabajo suene bien, así que exigen una atención a las propias palabras. En algunas dimensiones, al menos, esto es un ideal reconocido de la tradición filosófica, sobre todo de la filosofía analítica. No obstante, esta atención ha sido frecuentemente unilateral —con un solo oído, podría decirse—. Alentamos, correctamente, una preocupación porque *sea* verdadero (y exacto, etc.) lo que decimos, pero menos porque *suene* verdadero. Una buena pregunta, al menos para empezar, es si lo que uno ha escrito es algo que una persona adulta, interesada, inteligente podría decir a otra acerca de estos temas. Desde luego que no es la única pregunta, y no siempre se aplica, puesto que la filosofía no es simplemente conversación ordinaria. La filosofía es, más bien, en estos campos, la extensión de nuestras más serias preocupaciones por otros medios, pero al menos debería introducir nuestras preocupaciones ordinarias de una forma humanamente reconocible. De mucha filosofía que pretendidamente trata sobre temas éticos o políticos (y también de otros tipos), uno podría razonablemente preguntar: ¿qué pensaría si alguien que me hablara sonara realmente así?

Alguna filosofía moral puede asumir un tono diferente, más formal, por la razón especial de que está dirigida a posibles regulaciones. Es un lugar común que en las sociedades modernas muchos asuntos que solían ser preocupaciones privadas, o eran

conducidos en circunstancias privadas por profesionales, tales como practicantes médicos, son ahora conducidos en instituciones públicas, donde quienes forman parte de ellas deberían de ser públicamente responsables. Esto puede requerir que las decisiones sean gobernadas por principios declarados. Quienes discuten áreas tales como la ética médica, frecuentemente se preguntan qué sería una regulación pública justa, apropiada y funcional. Pero entonces *eso* es lo que están haciendo, y el estilo debería ser reconocido como una posibilidad para la filosofía, pero sólo una entre otras; una que apropiadamente usa un tono legal, regulativo o político, como dando a entender que ha habido un cambio en el orden público.

4

La exigencia de que la filosofía escuche lo que ella misma dice, conduce a las preocupaciones de Wittgenstein —particularmente, quizá, como las interpretaba Stanley Cavell—. [136] La posición de este pensador en relación con la filosofía analítica —y, por implicación, con la postanalítica— es anómala. Por un lado, era enteramente opuesto a la asimilación de la filosofía a la ciencia, y a la concepción misma de teoría filosófica (esto es tan verdadero del *Tractatus*, desde luego, como de su obra posterior). Por otro lado, él es seguramente, en términos del canon, una figura líder de la filosofía analítica.

Esto debería ser embarazoso, y el uso que se hace de su presencia en el canon muestra que lo es. En su mayor parte, Wittgenstein ha sido asimilado ahora a la empresa teórica, en la filosofía del lenguaje o la filosofía de la mente. Esto está bien en el nivel donde las ideas de cualquier filósofo se pueden convertir, de una forma más bien indeterminada, en parte de los recursos intelectuales de la materia, pero no es muy bueno para quienes lo tratan como parte del canon, lo cual implica leer algunos de sus libros, que (no importa qué estatus histórico o autoral se les atribuya) repetidamente dejan en claro que esto no era lo que debía suceder a estas ideas. La mayor parte de lo que dice Wittgenstein acerca de la naturaleza de la filosofía —su imposibilidad, en efecto— tiene que dejarse de lado como una excentricidad.

La obra de Wittgenstein no está bien protegida contra este tratamiento que se le ha dado. En parte, esto se debe a que las cuestiones de las que hablaba se hallan, desde luego, entre las preocupaciones centrales de la filosofía analítica, y a que las ideas a las que reaccionó fueron las de Frege y Russell, entre otros. Parcialmente, se debe a que casi ninguno de sus últimos escritos fue publicado por él, y, dado que en ellos se estaba hablando a sí mismo, no le ayuda al lector a darles el lugar adecuado. Y precisamente porque estaba hablándose a sí mismo, no pensó mucho en lo que sus lectores podrían hacer con estos pensamientos, así que no las preparó en contra de lo que hubiera considerado un mal uso.

En todos estos sentidos, vale la pena decirlo, difiere de Nietzsche, pues los arduos esfuerzos que se han hecho para asimilar a este último a las formas de filosofía estándar, en cualquier estilo, han sido una falla continua y saludable. Pero Nietzsche escribió sus

libros él mismo (excepto por uno que no existe); su tema se traslapa menos con las preocupaciones de la filosofía analítica; y tenía un sentido poderoso tanto de la importancia y originalidad de lo que estaba diciendo, como de la banalidad de los pensadores a los cuales podría ser asimilado, y minó o llenó de trampas sus obras para evitar que fueran asimiladas; así que es solamente ahora (como predijo) que estamos empezando a tener alguna idea de cuál es la mejor manera de leerlo.

Hay, desde luego, una corriente de opinión en la filosofía analítica que se resiste a esta asimilación de Wittgenstein a la empresa teórica, pero esto, de todos modos, suele revelar el bochorno del que hablábamos. No tiene nada muy interesante que decir acerca del caso de tal filósofo, ni de por qué vale la pena leer este canon. El modelo terapéutico de filosofía que es recuperado de algunos de los textos de Wittgenstein no es convincente, sobre todo porque, concedidos sus intereses, la enfermedad está representada como los intereses de la filosofía analítica misma, a la cual el curso está de otra manera invitando a la gente; uno está destinado a recordar la famosa observación de Karl Kraus de que el psicoanálisis es la enfermedad para la cual él mismo es la cura. Más aún, la idea es que la filosofía está separada de la conciencia cotidiana; su habla, de lo que alguien diría fuera de su campo, y la pregunta debe ser, entonces, cómo debe ser tomada. El modo en que no debe ser tomada —y éste es el punto más importante— sería una banalidad.

La idea de que debería tomarse como una banalidad fue un error de alguna filosofía de “lenguaje ordinario”, que identificaba dicho lenguaje ordinario con una vida convertida en ordinaria por la sustracción de la imaginación, y esto lo ha llevado a cabo alguien que, ahora, defiende a Wittgenstein en contra de la asimilación de su trabajo a la filosofía teórica. Esto es parcialmente —desde luego, pienso que sustancialmente— culpa del mismo Wittgenstein. Lo cotidiano y prefilosófico, de lo cual se supone que la filosofía está separada, se halla pobremente ubicada por él. En un bien conocido pasaje[137] escribió: “Y realmente obtiene una conexión extraña cuando el filósofo trata de explicarnos la relación entre nombre y cosa observando un objeto en frente de él y repitiendo un nombre (incluso la palabra “esto”) innumerables veces, pues los problemas filosóficos surgen cuando el lenguaje *se va de vacaciones*”. [*Wenn die Sprache feiert*: hay un problema con la traducción. *Feiern* puede significar “celebrar”, “tener un festival”; de hecho, “disfrutar de sí mismo”.][138] También puede significar algo así como “tomarlo con calma”. Entonces, ¿qué se supone que está haciendo el lenguaje? El contraste que Wittgenstein parece evocar es reminiscente de las líneas que abren el *Julio César* de Shakespeare:

¡Largo de aquí! A vuestras casas, gente ociosa, marchad a vuestras casas.
¿Es hoy día festivo? ¿Qué ignoráis
siendo artesanos, que no debéis
salir en día de trabajo sin los distintivos
de vuestra profesión?

Pero no puede ser simplemente que en el extraño trabajo del filósofo el lenguaje no cumpla su función ordinaria. La labor del filósofo está separada de la vida ordinaria, pero

una festividad pública, después de todo, es parte de la vida ordinaria. No puede significar que para el lenguaje no haya un juego legítimo (es, finalmente, Wittgenstein quien caracteriza el uso ordinario del lenguaje como “juego del lenguaje”).

Podría significar que el lenguaje está haciendo una cosa (algo no muy serio) cuando debería estar haciendo otra, pero, ¿qué “debería” es ese? (¿Cómo sabemos que las palabras son mecánicas, y que es un día laboral?) Después de todo, sería posible que alguien hiciera justamente lo que el filósofo está haciendo, y que lo hiciera para algún propósito humano reconocible y no alienado —por ejemplo si, quizá, fuera Petrarca diciendo el nombre “Laura” muchas veces ante una estatua de dicha mujer—. Si excluyéramos una actividad como esa por hacer un uso irresponsable del lenguaje, estaríamos aplicando un canon de lo ordinario que nos regresaría a la filistea banalidad de la “filosofía del lenguaje ordinario”. Más bien, tendremos que decir: el desempeño del filósofo está separado del mundo (un mundo del cual la imaginada actividad de Petrarca, por ejemplo, podría ser parte) porque somos introducidos más efectivamente a su supuesto propósito —encontrar la esencia de la relación nominal— a través de un uso del lenguaje que no tiene en absoluto ni un propósito ni una importancia humanos independientemente reconocibles.

La famosa frase es equívoca porque no es verdad que observando una festividad, celebrando o pasándola bien, el lenguaje deje de encarnar los propósitos humanos. Es mucho menos verdadero que el lenguaje esté etiquetado por adelantado con los signos de su profesión; que podamos enlistar qué usos del lenguaje son no alienados y, en ese sentido, serios, y entonces reconocer la filosofía por contraste con ellos. Más bien, de cualquier uso podemos preguntar “¿qué está haciendo el lenguaje aquí?”, y a veces obtendremos una buena respuesta, que es al mismo tiempo una en la que no hubiéramos pensado. La respuesta del filósofo, cuando es del tipo ilustrado por el ejemplo de Wittgenstein, no es buena, porque (como supone este pensador) la mejor expresión de la perplejidad del filósofo ha de encontrarse en usos del lenguaje que nuevamente plantean la pregunta “¿qué está haciendo el lenguaje aquí?”, eso no puede responderse correctamente.

Esto conduce a un punto importante. En el caso que menciona Wittgenstein, como en la mayoría —quizá todos— de los casos que le preocupaban, los usos del lenguaje que mejor expresan la perplejidad del filósofo se hallan separados de *todo* propósito humano, así que podrían ser razonablemente considerados carentes de significado.[139] Están, por así decirlo, atemporalmente fuera de lugar, porque no encajan con ningún propósito humano concebible, excepto el extraviado impulso filosófico que supuestamente ilustran. Es debido a que las compulsiones filosóficas de este tipo son expresadas por medio de un lenguaje que está separado de *todo*, que Cavell interpreta esta línea de pensamiento de Wittgenstein en términos metafísicos. En su análisis de este último, Cavell enfatiza lo que él considera las posibilidades omnipresentes del escepticismo, implícitas en la condición humana, y las relaciones de esas posibilidades con lo que, erróneamente, tiende a pensar que se halla más allá de los límites de nuestra vida. Esto, desde luego, supone que lo filosóficamente separado es de alguna manera ininteligible. Ésta es una idea kantiana, y

retiene la conexión de la filosofía con lo atemporal y lo universal, pero hay otros modos en los que la filosofía puede ser separada del habla humana, y Wittgenstein, quien en su obra posterior lucha precisamente en contra del tipo de filosofía que había creado anteriormente, que aspiraba a la más completa generalidad, no nos ayuda a ver esto.

El lenguaje que proviene de los filósofos, y que nos impacta por ser ofensivo o inapropiado cuando lo escuchamos propiamente, puede muy bien no ser *ininteligible*; puede ser más bien falso, mecánico, desapegado o cursi. Éstos, significativamente, son términos para valorar los escritos de ficción o de otras artes, no términos de diagnóstico semántico.

Es verdad que el propio proyecto de Cavell, a diferencia del de Wittgenstein, frecuentemente parece relacionarse con la historia o con nuestra situación cultural presente, particularmente a través de su compromiso con el modernismo. Pero, concedidas sus presuposiciones generales, en alguna parte derivadas de la práctica de Wittgenstein, este compromiso es él mismo conducido en términos metafísicos, y esto deja espacio para no mucho más que el pensamiento de que hay algo especialmente acerca del mundo moderno que es metafísicamente alienador. Esto entonces tiende a imponer una imagen de una caída de la unidad prima, y en esta etapa, no es sorprendente que se escuchen resonancias heideggerianas.

5

Pero, ¿por qué no deberíamos, al demandar que se ponga atención a la forma en que suena la filosofía, introducir *genuinas* categorías históricas? Lo que podemos honesta y útilmente decir, que pueda sonar verdadero, es en parte una función de nuestro tiempo, y demanda que aprehendamos con la imaginación lo que está vivo y lo que está muriendo en nuestro tiempo. Si esto es en parte un asunto de historia (de historia real, y no de una mítica historia de separación), es aún más probable que sólo pueda ser transmitido empleando recursos más imaginativos y expresivos de lo que permite o alienta la versión estándar de la filosofía analítica.

Deberíamos aceptar la idea de que algo de lo que ha preocupado a la filosofía, e indudablemente le preocupa, es la cuestión de lo que es inequívocamente algún tipo de historia, una idea que en la cultura de la filosofía británica sólo ha sido tomada seriamente por Collingwood y por Berlín. La así llamada esencia de cierto valor (para tomar particularmente el caso de los valores) puede ser tan esquemática o indeterminada que sólo puede entenderse con referencia a formaciones históricas particulares. Nada que tenga historia puede ser definido, como correctamente dijera Nietzsche, y nuestras virtudes y nuestros valores ciertamente tienen historia.[140]

En un libro reciente, Mary Warnock argumentó que no podía haber algo como una filosofía feminista, porque la filosofía se ocupaba solamente de lo humanamente universal.[141] Puede responderse más bien rápidamente a este argumento, y en términos similares a los suyos: las prácticas y actitudes de las que se ocupa la filosofía feminista

forman parte de lo humanamente universal. Pero lo importante es que el principio del argumento de Mary Warnock debe ser enteramente erróneo. Ya sabemos qué cosas nos ocupan cuando abordamos las cuestiones de la filosofía política y moral, y ¿cómo podemos conocer los modos en que esas cuestiones se relacionan con lo humanamente universal hasta que las miramos y vemos?

Quiero enfatizar que la imagen que estoy tratando de bosquejar no se deriva de la filosofía externa. Preferiría, con mucho, que retuviéramos la categoría de filosofía y nos situáramos dentro de ella, en vez de pretender que una investigación que aborde estos asuntos con recursos más ricos e imaginativos represente “el fin de la filosofía”. Las tradiciones de la filosofía exigen que reflexionemos sobre las presuposiciones de lo que pensamos y sentimos. La aseveración que ahora estoy haciendo desde aquí, desde dentro de la materia, es que en ciertas áreas, al menos, esta misma exigencia no puede ser adecuadamente satisfecha a menos que vayamos más allá de las concepciones del modo correcto de proceder que se hallan demasiado estrechamente asociadas con los modelos inexpresivos extraídos, quizá inconscientemente, de las ciencias.

He sugerido que esto es así al menos por dos razones. La misma estructura de nuestros valores, primero, puede ser de tal modo que requiera un entendimiento histórico, que incluya una aprehensión imaginativa del pasado. Segundo, la filosofía no estará hablando de nuestras preocupaciones a menos que suene bien, a menos que la forma del trabajo mismo exprese lo que el escritor siente que está viviendo, o de lo contrario resultará derivado y falso, y es probable que eso sea un logro imaginativo. No conseguir eso —que es requerido por las metas tradicionales de la filosofía misma— muy bien puede provocar las fallas que atraen las críticas aplicables a la literatura imaginativa, no solamente las que son familiares a la filosofía.

Alguien puede, desde luego, como están las cosas, decir que una parte de la filosofía es inimaginativa, pero por lo regular esto significa que hay alguna línea de argumento que no ha sido explorada, alguna teoría que podría haber sido aplicada, alguna objeción que podría haberse dejado de lado. Pero deberíamos recordar que un trabajo puede ser no imaginativo no porque esté mal argumentado, sino por presentarse a las personas equivocadas; no porque haya perdido un debate, sino por falta de comprensión histórica y psicológica; no porque deje de ser inteligente, sino porque es tonto.

Podemos soñar en una filosofía que sería plenamente verdadera y honestamente útil. Eso que requiere, desde luego, una combinación imposible de características.

Sus argumentos serían ordenados —eso sería parte de lo que la hizo filosofía—, pero retendría la posibilidad, tanto en su contenido como en su forma, de ser perturbadora en vez de tranquilizadora. Le hablaría al lector maduro en términos que éste pudiera reconocer como dignos de ser escuchados. Necesitaría los recursos de una imaginación expresiva para hacer casi cualquiera de las cosas que necesitaba hacer, pero sería más probable que motivara el asentimiento la mayor parte del tiempo por la retórica de la proposición clara más que por una asociación imaginativa manifiesta. Esto no se debe a que la filosofía no sea literatura. Alguna filosofía necesita virtudes literarias y también debería temer el fracaso en tanto literatura. Incluso puede ser a veces literatura, pero sólo

cierto tipo de literatura.

Tendría que ser fiel a la historia de la filosofía, o al menos a algo de ella. No es una vacía perogrullada decir que la filosofía es aquello de lo que es historia la historia de la filosofía. Significa que la filosofía es una empresa que ha logrado ciertas cosas, y que se encuentra en cierto estado. La filosofía necesita permanecer fiel a su práctica, o al menos tener claro por qué está rechazando partes de ella. Hay, desde luego, deformaciones profesionales, y son, desde luego, severas, pero la desconfianza en ellas debería ser en sí misma desconfianza profesional, podríamos decir, stravinskiana.

La filosofía debe sostener firmemente la idea de que su fuerza para confrontar nuestros sufrimientos yace en lo que puede especialmente traerles, y lo que puede traer conlleva su dificultad inevitable. Sus poderes especiales implican que lo que dice jamás es obvio. Hay más de un modo de hacer eso, y solamente algunos de ellos son técnicos (piensen en el último Wittgenstein, o en Nietzsche). Hay muchas cosas verdaderas y obvias que decir de los horrores del mundo, y muchos tipos de escritura pueden y deben decir dichas cosas de un modo obvio; pero éstas no son por lo regular las cosas que la filosofía, si es que ha de ser útil en sus especiales modos, tiene razones para decir; o si en una ocasión tiene razón para hacerlo, lo que no será obvio será su razón para decirlas.

Este énfasis no es simplemente el producto de un prejuicio modernista de mi parte. Concedidas todas las cosas que la filosofía tiene que ordenar, como su uso del argumento, el reconocimiento de que no debe ser obvia solamente registra la consecuencia para la filosofía del requerimiento general de que al escribir, si es que ha de ser verdadera, debe escuchar lo que está diciendo.

Bernard Williams

Publicaciones filosóficas completas

Williams también escribió muchos ensayos sobre ópera y muchos artículos de periodismo ocasional. Los primeros serán publicados todos juntos en un libro a su debido tiempo. Los segundos, muchos de los cuales son piezas filosóficas, están incluidos en una selecta bibliografía intitulada “Bernard Williams: Writings of Political Interest” en *In the Beginning Was the Deed* (véase la referencia abajo).

Libros

Morality: An Introduction to Ethics, Nueva York, Harper & Row, 1972, Harmondsworth, Penguin, 1973, Cambridge, Cambridge University Press, 1976; Canto edition, con nueva introducción, 1993. Traducción alemana: *Der Begriff der Moral*, Leipzig, Reclam, 1978. Traducción rumana: *Introducere in etica*, Bucarest, Editura Alternative, 1993. Traducción francesa: véase *La fortune morale* (1994). Traducción italiana (de Canto edition): *La moralità: un'introduzione all'etica*, Roma, Einaudi, 2000. Traducción polaca: *Moralność: Wprowadzenie do etyki*, Warsaw, Fundacia Aletheia, 2000.

Problems of the Self: Philosophical Papers 1956-1972, Cambridge, Cambridge University Press, 1973. Traducción alemana: *Probleme des Selbst*, Leipzig, Reclam, 1978. Traducción en español: *Problemas del yo*, México, Universidad Nacional Autónoma de México, 1986. Traducción italiana: *Problemi dell'io*, Milán, Il Saggiatore, 1990.

A Critique of Utilitarianism, en *Utilitarianism: For and Against*, por J.J.C. Smart y Bernard Williams, Cambridge, Cambridge University Press, 1973. Traducción alemana: *Kritik des Utilitarismus*, Frankfurt am Main, Klostermann, 1979. Traducción en español: *Utilitarismo: pro y contra*, Madrid, Tecnos, 1981. Traducción italiana: *Utilitarismo: un confronto*, Nápoles, Bibliopolis, 1985. Traducción francesa: *Utilitarisme: le pour et le contre*, Le Champ Ethique, Ginebra, Labor et Fides, 1997.

Descartes: The Project of Pure Enquiry. Harmondsworth: Penguin; Hassocks: Harvester

- Press, 1978. Reimpreso con una nueva introducción por John Cottingham, Londres, Routledge, 2005. Traducción alemana: *Descartes: Das Vorhaben der reinen philosophischen Untersuchung*, Frankfurt am Main, Athenaem, 1981. Traducción en español: *Descartes: el proyecto de la investigación pura*, México, Universidad Nacional Autónoma de México, 1995.
- Moral Luck: Philosophical Papers 1973-1980*, Cambridge, Cambridge University Press, 1981. Traducción alemana: *Moralischer Zufall*, Königstein, Hain, 1984. Traducción italiana: *Sorte morale*, Milán, Il Saggiatore-Mondadori, 1987. Traducción en español: *La fortuna moral*, México, Universidad Nacional Autónoma de México, 1993.
- Ethics and the Limits of Philosophy*. Londres: Fontana Books; Cambridge: Harvard University Press, 1985. Traducción italiana: *L'etica e i limiti della filosofia*, Bari, Laterza, 1987. Traducción francesa: *L'éthique et les limites de la philosophie*, París, Editions Gallimard, 1990. Traducción japonesa: *Ikikata ni tsuite Tetsugaku wa Nani ga Ieru ka*, Tokio, Sangyoutosho K.K., 1993. Traducción en español: *La ética y los límites de la filosofía*, Caracas, Monte Ávila Editores, 1997. Traducción alemana: *Ethik und die Grenzen der Philosophie*, Hamburgo, Rotbuch Verlag, 1999.
- Shame and Necessity*, Sather Classical Lectures, vol. 57, Berkeley y Los Ángeles, University of California Press, 1993. Traducción francesa: *La honte et la nécessité*, Presses Universitaires de France, 1997. Traducción alemana: *Scham, Schuld und Notwendigkeit*, Polis, vol. 1, Berlín, Akademie Verlag, 2000.
- La fortune morale*, París, Presses Universitaires de France, 1994. (= *Morality* + selecciones de *Problems of the Self*, *Moral Luck*, y *Making Sense of Humanity*.)
- Making Sense of Humanity and Other Philosophical Papers 1982-1993*, Cambridge: Cambridge University Press, 1995.
- Der Wert der Wahrheit*, Traducido por Joachim Schulte, Viena, Passagen Verlag, 1998.
- Plato: The Invention of Philosophy*, Londres, Phoenix/Orion, 1998, reimpreso en *The Sense of the Past* (2006), detalles abajo.
- Ile wolności powinna mieć wola?*, Warsaw, Fundacja Aletheia, 1999. (= Traducción polaca de artículos seleccionados.)
- Truth and Truthfulness: An Essay in Genealogy*, Princeton, Princeton University Press, 2002. Traducción alemana: *Wahrheit und Wahrhaftigkeit*, Frankfurt am Main, Surhrkamp, 2003. Traducción italiana: *Genealogia della verità: storia e virtù del dire il vero*, Roma, Fazi Editore, 2005. Traducción francesa: Título aún desconocido, París, Editions Gallimard, 2005.
- In the Beginning Was the Deed: Realism and Moralism in Political Argument*, editado por Geoffrey Hawthorn, Princeton, Princeton University Press, 2005.
- The Sense of the Past: Essays in the History of Philosophy*, editado por Myles

Burnyeat, Princeton, Princeton University Press, 2006.

Philosophy as a Humanistic Discipline, editado por A. W. Moore, Princeton, Princeton University Press, 2006.

Libros editados

(Con A.C. Montefiore.) *British Analytical Philosophy*, Londres, Routledge, 1966. Traducción italiana: *Filosofia analitica inglese*, Roma, Lerici 1967.

(Con A. K. Sen.) *Utilitarianism and Beyond*, Cambridge, Cambridge University Press, 1982. Traducción italiana: *Utilitarismo e oltre*, Milán, Il Saggiatore, 1984.

Artículos

En general, no se incluyen las referencias a reimpressiones de artículos, excepto en:

[PS]: *Problems of the Self*

[ML]: *Moral Luck*

[MSH]: *Making Sense of Humanity*

[IBD]: *In the Beginning Was the Deed*

[SP]: *The Sense of the Past*

[PHD]: *Philosophy as a Humanistic Discipline*

“Tertullian’s Paradox”, en *New Essays in Philosophical Theology*, editado por Anthony Flew y Alasdair Macintyre, Londres, SCM Press, 1955. [PHD]

“Personal Identity and Individuation”, *Proceedings of the Aristotelian Society* 57 (1956-1957). [PS]

“Metaphysical Arguments”, en *The Nature of Metaphysics*, editado por D. F. Pears, Londres, Macmillan, 1957. [PHD]

“Pleasure and Belief”, *Proceedings of the Aristotelian Society*, suppl. vol. 33 (1959). [PHD]

“Descartes”, en *A Dictionary of Philosophy and Philosophers*, editado por J. O. Urmsen, Londres, Hutchinson, 1960.

“Personal Identity and Bodily Continuity—a Reply”, *Analysis* 21 (1960). [PS]

- “Mr Strawson on Individuals”, *Philosophy* 36 (1961). [PS]
- “The Individual Reason”, *The Listener*, noviembre 16, 1961.
- “Democracy and Ideology”, *Political Quarterly* 32 (1961).
- “The Idea of Equality”, en *Politics, Philosophy and Society*, editado por Peter Laslett y W. G. Runciman, Oxford, Blackwell, 1962. [PS y IBD]
- “Aristotle on the Good: A Formal Sketch”, *Philosophical Quarterly* 12 (1962). [SP]
- “La certitude du cogito”, en *Cahiers du Royaumont*, vol. 4, París, Editions de Minuit, 1962.
- Texto en inglés: “The Certainty of the Cogito”, en *Descartes: A Collection of Critical Essays*, editado por W. Doney, Nueva York, Doubleday, 1967.
- “Freedom and the Will”, con un *postscript*, en *Freedom and the Will*, editado por D. F. Pears, Londres, Macmillan, 1963.
- “Imperative Inference”, *Analysis*, suppl. vol. 23 (1963). [PS]
- “Hume on Religion”, en *David Hume: A Symposium*, editado por D. F. Pears, Londres, Macmillan, 1963. [SP]
- “Ethical Consistency”, *Proceedings of the Aristotelian Society*, suppl. vol. 39 (1965). [PS]
- “Morality and the Emotions”, discurso inaugural, Bedford College, Londres, 1965, reimpresso en *Morality and Moral Reasoning: Five Essays in Ethics*, editado por J. Casey, Londres, Methuen, 1971. [PS]
- “Imagination and the Self”, *British Academy Annual Philosophical Lecture* [Conferencia filosófica anual de la Academia Británica], 1966. [PS]
- “Consistency and Realism”, *Proceedings of the Aristotelian Society*, suppl. vol. 40 (1966). [PS]
- “Descartes”, en *The Encyclopedia of Philosophy*, editado por P. Edwards. Nueva York: Macmillan y Free Press, Londres, Collier-Macmillan, 1967.
- “Hampshire, S. N”, en *ibid.*
- “Rationalism”, en *ibid.*
- “Knowledge and Meaning in the Philosophy of Mind”, *Philosophical Review* 79 (1968). [PS]
- “Has ‘God’ a Meaning?”, *Question* 1 (1968). Traducción alemana: “Der unverzichtbare Gehalt des christlichen Glaubens”, en *Glaube und Vernunft*, editado por N. Hoerster, Munich, Deutsche Taschenbuch Verlag, 1979, Leipzig, Reclam, 1985.
- “Descartes’ Ontological Argument: A Comment”, en *ibid*, *Fact and Existence*, editado

- por J. Margolis, Oxford, Blackwell, 1969.
- “Existence-Assumptions in Practical Thinking: Reply to Körner”, en *ibid.*
- “Philosophy”, en *General Education: A Symposium on the Teaching of Non-Specialists*, editado por Michael Yudkin, Harmondsworth, Penguin, 1969.
- “The Self and the Future”, *Philosophical Review* 79 (1970). [PS]
- “Are Persons Bodies?”, en *The Philosophy of the Body: Rejections of Cartesian Dualism*, editado por S. Spicker, Chicago, Quadrangle Books, 1970. [PS]
- “Genetics and Moral Responsibility”, en *Morals and Medicine*, Londres, BBC Publications, 1970.
- “Deciding to Believe”, en *Language, Belief, and Metaphysics*, editado por H. E. Kiefer y M. K. Munitz. Albany, State University of New York Press, 1970. [PS]
- “The Temporal Ordering of Perceptions and Reactions: Reply to O’Shaughnessy”, en *Perception: A Philosophical Symposium*, editado por F. Sibley, Londres, Methuen, 1971.
- “Conversation on Moral Philosophy”, en *Modern British Philosophy*, editado por B. Magee, Londres, Secker and Warburg, 1971.
- “Knowledge and Reasons”, en *Problems in the Theory of Knowledge*, editado por G. H. von Wright, The Hague, Martinus Nijhoff, 1972. [PHD]
- “The Analogy of City and Soul in Plato’s Republic”, en E. N. Lee, A. P. D. Mourelatos y R. M. Rorty (eds.), *Exegesis and Argument: Studies in Greek Philosophy. Essays Presented to Gregory Vlastos*, The Hague, Martinus Nijhoff, 1973. [SP]
- “The Makropoulos Case: Reflections on the Tedium of Immortality”. [PS]
- “Egoism and Altruism”. [PS]
- Observaciones en *The Law and Ethics of AIDS and Embryo Transfer*, CIBA Foundation Symposium 17 / Simposio 17 de la Fundación CIBA,. Amsterdam, Elsevier/North-Holland, 1973.
- “Wittgenstein and Idealism”, en *Understanding Wittgenstein*, editado por Godfrey Vesey, Royal Institute of Philosophy Lectures 7, Londres, Macmillan, 1974. [ML y SP]
- “The Truth in Relativism”, *Proceedings of the Aristotelian Society* 75 (1974-1975). [ML]
- “Rawls and Pascal’s Wager”, *Cambridge Review*, febrero 28, 1975. [ML]
- “Persons, Character and Morality”, en *The Identities of Persons*, editado por A. Rorty. Berkeley y Los Ángeles, University of California Press, 1976. [ML]
- “Utilitarianism and Moral Self-Indulgence”, en *Contemporary British Philosophy*,

- editado por H. D. Lewis, Ser. 4, Londres, Allen & Unwin, 1976. [ML]
- “Moral Luck”, *Proceedings of the Aristotelian Society*, suppl. vol. 59 (1976). [ML]
- “The Moral View of Politics”, *The Listener*, junio 3, 1976.
- “Thinking about Abortion”, *The Listener*, septiembre 1, 1977.
- “Linguistic Philosophy”, en *Men of Ideas*, editado por B. Magee, Londres, BBC Publications, 1978.
- “Politics and Moral Character”, en *Public and Private Morality*, editado por Stuart Hampshire, Cambridge, Cambridge University Press, 1978. [ML]
- Introducción a Isaiah Berlin, *Concepts and Categories: Philosophical Essays*, Londres, Hogarth Press, 1978.
- Conclusión a *Morality as a Biological Phenomenon*, editado por Gunther Stent, Berlín, Dahlem Konferenzen, 1978, Berkeley y Los Ángeles, University of California Press, 1980.
- [Con David Wiggins]. Introducción a *Ethics, Value and Reality: Selected Papers of Aurel Kolnai*, Londres, Athlone, 1978.
- “Another Time, Another Place, Another Person”, en *Perception and Identity: Essays Presented to A. J. Ayer; with His Replies to Them*, editado por G. Macdonald, Londres, Macmillan, 1979. [ML]
- “Conflicts of Values”, en *The Idea of Freedom: Essays in Honor of Isaiah Berlin*, editado por A. Ryan, Oxford, Oxford University Press, 1979. [ML]
- “Internal and External Reasons”, en *Rational Action: Studies in Philosophy and Social Science*, edited by T. R. Harrison. Cambridge, Cambridge University Press, 1979. [ML]
- “Political Philosophy and the Analytical Tradition”, en *Political Theory and Political Education*, editado por M. Richter, Princeton, Princeton University Press, 1980. [PHD]
- “Moral Obligation and the Semantics of ‘Ought’”, *Proceedings of the Fifth Kirchberg Wittgenstein Conference*, 1980, versión revisada, “‘Ought’ and Moral Obligation”. [ML]
- “L’éthique et la philosophie analytique”, *Critique* agosto-septiembre, París, 1980.
- “Philosophy”, en *The Legacy of Greece: A New Appraisal*, editado por M. Finley, Oxford, Oxford University Press, 1981. [SP]
- “Justice as a Virtue”, en *Essays on Aristotle’s Ethics*, editado por A. Rorty, Berkeley y Los Ángeles, University of California Press, 1981. [ML y SP]

- “Practical Necessity”, en *Philosophical Frontiers of Christian Theology: Essays Presented to Donald Mackinnon*, editado por S. Sutherland y B. Hebblethwaite, Cambridge, Cambridge University Press, 1982. [ML]
- “Cratylus’ Theory of Names and Its Refutation”, en *Language and Logos: Studies in Ancient Greek Philosophy Presented to G.E.L. Owen*, editado por M. Schofield y M. Nussbaum, Cambridge, Cambridge University Press, 1982. [SP]
- “The Point of View of the Universe: Sidgwick and the Ambitions of Ethics”, Henry Sidgwick Memorial Lecture [Conferencia conmemorativa Henry Sidgwick], 1982, *Cambridge Review*, mayo 7, 1982. [MSH y SP]
- “Evolution, Ethics, and the Representation Problem”, en *Evolution from Molecules to Men*, editado por D. S. Bendall, Cambridge, Cambridge University Press, 1983. [MSH]
- “Space Talk: The Conversation Continued” (comentario sobre *Social Justice in the Liberal State* de B. Ackerman), *Ethics* 9.3 (1983).
- “Descartes’ Use of Skepticism”, en *The Skeptical Tradition*, editado por Myles Burnyeat, Berkeley y Los Ángeles, University of California Press, 1983. [SP]
- “Professional Morality and Its Dispositions”, en *The Good Lawyer: Lawyers’ Roles and Lawyers’ Ethics*, editado por David Luban, *Maryland Studies in Public Philosophy* [Estudios Maryland en filosofía pública], Totowa, NJ, Rowman and Allenheld, 1983. [MSH]
- “Präsuppositionen der Moralität”, en *Bedingungen der Möglichkeit: “Transcendental Arguments” und Transcendentales Denken*, editado por E. Schaper y W. Vossenkuhl, Stuttgart, Klett-Cotta, 1984.
- “The Scientific and the Ethical”, en *Objectivity and Cultural Divergence*, editado por S. Brown, *Royal Institute of Philosophy Lectures* 17, Cambridge, Cambridge University Press, 1984. [Versión abreviada del capítulo 8 de *Ethics and the Limits of Philosophy*.]
- “Morality, Scepticism and the Nuclear Arms Race”, en *Objections to Nuclear Defence: Philosophers on Deterrence*, editado por N. Blake and K. Pole, Londres, Routledge, 1984.
- “Formal and Substantial Individualism”, *Proceedings of the Aristotelian Society* 85 (1984-85) [MSH].
- “Theories of Social Justice-Where Next?”, en *Equality and Discrimination: Essays in Freedom and Justice*, editado por S. Guest y A. Milne, *Archiv für Rechtsund Sozialphilosophie* 21, Stuttgart: F. Steiner, 1985.
- “Ethics and the Fabric of the World”, en *Morality and Objectivity: A Tribute to J. L. Mackie*, editado por T. Honderich, Londres, Routledge, 1985. [MSH]

- “What Slopes Are Slippery?”, en *Moral Dilemmas in Modern Medicine*, editado por M. Lockwood, Oxford, Oxford University Press, 1985. [MSH]
- How Free Does the Will Need to Be? The Lindley Lecture* [La conferencia Lindley] Lindley, 1985, Lawrence, University of Kansas Press, 1986. [MSH]
- “L’entrevista di *Politeia*: Bernard Williams”, *Politeia* (Milán) (Invierno 1986).
- “Formalism and Natural Language in Moral Philosophy”, en *Mérites et limites des méthodes logiques en philosophie*, editado por Jules Vuilleman, París, J. Vrin, para Fondation Singer-Polignac, 1986.
- “Types of Moral Argument against Embryo Research”, en *Human Embryo Research: Yes or No?*, Londres, Tavistock Publications, para la fundación CIBA, 1986, Reimpreso en *BioEssays* 6 (1987).
- “Hylomorphism”, *Oxford Studies in Ancient Philosophy* 4 (1986), *Festschrift* para J. L. Ackrill, editado por Michael Woods. [SP]
- Introducción a René Descartes, *Meditations on First Philosophy with Selections from the Objections and Replies*, editado por J. Cottingham, Cambridge, Cambridge University Press, 1986. [SP]
- “Reply to Simon Blackburn” (Reseña de *Ethics and the Limits of Philosophy*), *Philosophical Books* 27 (1986).
- Comments on Amartya Sen’s Tanner Lectures, 1985, en Amartya Sen, *The Standard of Living*, Cambridge, Cambridge University Press, 1987.
- “The Primacy of Dispositions”, en *Education and Values: The Richard Peters Lectures*, editado por Graham Haydon, Londres, University of London Institute of Education, 1987. [PHD]
- “Descartes”, en *The Great Philosophers: An Introduction to Western Philosophy*, editado por B. Magee, Londres, BBC Publications, 1987.
- “The Structure of Hare’s Theory”, en *Hare and Critics: Essays in Moral Thinking*, editado por Douglas Seanor y N. Fotion. *Festschrift* para R. M. Bare, Oxford, Oxford University Press, 1988. [PHD]
- “Formal Structures and Social Reality”, en *Trust: Making and Breaking Co-operative Relations*, editado por Diego Gambetta, Oxford, Blackwell, 1988. [MSH]
- “What Does Intuitionism Imply?”, en *Human Agency: Language, Duty, and Value. Philosophical Essays in Honor of J. O. Urmson*, editado por J. Moravcsik y C.C.W. Taylor, Stanford, Stanford University Press, 1988. [MSH]
- “Evolutionary Theory: Epistemology and Ethics”, en *Evolution and Its Influence*, editado por Alan Grafen, Herbert Spencer Lectures, 1986, Oxford, Oxford University Press,

1989. [MSH]
- “Dworkin on Community and Critical Interests”, *California Law Review* 77 (1989),
- “Modernita evita etica”, en *Etica e vita quotidiana*, Bolonia, Biblioteca del Mulino, 1989.
- Texto en inglés, “Modernity and the Substance of Ethical Life”. [IBD]
- “Voluntary Acts and Responsible Agents”, *Hart Lecture* [Conferencia Hart], Oxford, 1987, *Oxford Journal of Legal Studies* 10 (1989). [MSH]
- “Social Justice: The Agenda in Social Philosophy for the Nineties”, *Journal of Social Philosophy* 20 (1989).
- “Internal Reasons and the Obscurity of Blame”, *Logos* 10 (1989). [MSH] Réplica al Presidente. *Proceedings of the Aristotelian Society* 90 (1989-90).
- “Who Might I Have Been?”, en *Human Genetic Information: Science, Law and Ethics*, CIBA Foundation Symposium 149 / Simposio 149 de la Fundación CIBA, Chichester, John Wiley & Sons, 1990, versión revisada: “Resenting One’s Own Existence”. [MSH]
- “Notre vie éthique” (extracto del prefacio de la traducción francesa de *Ethics and the Limits of Philosophy*), *Esprit*, noviembre 11, 1990.
- “The Need to Be Sceptical”, *Times Literary Supplement*, febrero 16, 1990.
- “Making Sense of Humanity”, en *The Boundaries of Humanity: Humans, Animals, Machines*, editado por James Sheenan y Morton Sosna, Proceedings of Stanford University Centennial Conference / Actas de la Conferencia Centenaria de la Universidad de Stanford, 1987, Berkeley y Los Ángeles, University of California Press, 1991. [MSH]
- “Saint-Just’s Illusion: Interpretation and the Powers of Philosophy”, *London Review of Books*, agosto 29, 1991. [MSH]
- “Subjectivism and Toleration”, en *A. J. Ayer: Memorial Essays*, editado por A. Phillips Griffiths, Berkeley y Los Ángeles, University of California Press, 1992. [PHD]
- “Must a Concern for the Environment Be Centred on Human Beings?”, en *Ethics and the Environment*, editado por C.C.W. Taylor. Oxford, Corpus Christi College, 1992. [MSH]
- “Pluralism, Community and Left Wittgensteinianism”, *Common Knowledge* 1 (1992). [IBD]
- Introducción al *Theaetetus* de Platón, traducido por M. J. Levett, revisado por Myles Burnyeat, Indianápolis, Hackett, 1992. [SP]
- “Moral Incapacity”, *Proceedings of the Aristotelian Society* 92 (1992-93). [MSH]

- “Nietzsche’s Minimalist Moral Psychology”, *European Journal of Philosophy* 1 (1993). [MSH y SP]
- “Who Needs Ethical Knowledge?”, en *Ethics*, editado por A. P. Griffiths, Royal Institute of Philosophy Lectures, 1992, Cambridge, Cambridge University Press, 1993. [MSH]
- “Moral Luck: A Postscript”, en *Moral Luck*, editado por Daniel Statman, Albany, State University of New York Press, 1993. [MSH]
- “Les vertus de la vérité”, en *Le respect*, editado por C. Audard, París, Editions Autrement, 1993.
- “Pagan Justice and Christian Love”, en *Virtue, Love and Form: Essays in Memory of Gregory Vlastos*, editado por Terence Irwin and Martha C. Nussbaum, Edmonton, Academic Printing and Publishing, 1994. [SP]
- “Descartes and the Historiography of Philosophy”, en *Reason, Will and Sensation: Studies in Descartes’s Metaphysics*, editado por John Cottingham, Oxford, Oxford University Press, 1994. [SP]
- “The Actus Reus of Dr Caligari”, *Pennsylvania Law Review* 142 (mayo 1994). [PHD]
- “Replies”, en *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, editado por J.E.J. Altham and Ross Harrison, Cambridge, Cambridge University Press, 1995.
- “Ethics”, en *Philosophy: A Guide through the Subject*, editado por A. C. Grayling, Oxford, Oxford University Press, 1995.
- “Identity and Identities”, en *Identity: Essays Based on Herbert Spencer Lectures Given in the University of Oxford*, editado por Henry Harris, Oxford, Oxford University Press, 1995. [PHD]
- “Acts and Omissions, Doing and Not Doing”, en *Virtues and Reasons: Philippa Foot and Moral Theory. Essays in Honour of Philippa Foot*, editado por Rosalind Hursthouse, Gavin Lawrence y Warren Quinn, Oxford, Oxford University Press, 1995. [MSH]
- “La philosophie devant l’ignorance”, *Diogène* 169 (París, Editions Gallimard, 1995). Texto en inglés: “Philosophy and the Understanding of Ignorance”, en la correspondiente edición en inglés de *Diogenes* (Oxford, Berghahn Books). [PHD]
- “Acting as the Virtuous Person Acts”, en *Aristotle and Moral Realism*, editado por Robert Heinaman, Keeling Colloquium 1, 1994, Londres, UCL, 1995. [SP]
- “A Further Introduction”, en *The Blackwell Companion to Philosophy*, editado por N. Bunnin y E. P. Tsui-James, Oxford, Blackwell, 1996.
- “Truth in Ethics”, en *Truth in Ethics*, editado por Brad Hooker, Oxford, Blackwell,

1996.

- “Censorship in a Borderless World”, Faculty Lecture 16, Facultad de Artes y Ciencias Sociales, Singapur, Universidad Nacional de Singapur, 1996.
- “Values, Reasons, and the Theory of Persuasion”, en *Ethics, Rationality and Economic Behaviour*, editado por Francesco Farina, Frank Hahn y Stefano Van- nucci, Oxford, Oxford University Press, 1996. [PHD]
- “The Politics of Trust”, en *The Geography of Identity*, editado por Patricia Yaeger, Ann Arbor, University of Michigan Press, 1996.
- “Toleration: An Impossible Virtue?”, en *Toleration: An Elusive Virtue*, editado por David Heyd, Princeton, Princeton University Press, 1996.
- “History, Morality and the Test of Reflection”, en Christine Korsgaard with othhers, *The Sources of Normativity*, editado por Onora O’Neill, Cambridge, Cambridge University Press, 1996.
- “Truth, Politics and Self-Deception”, *Social Research* 63 (1996). [BD]
- “The Women of Trachis: Fictions, Pessimism, Ethics”, en *The Greeks and Us: Essays in Honor of Arthur W. H. Adkins*, editado por R. B. Loudon y P. Schollmeier, Chicago, University of Chicago Press, 1996. [SP]
- “La tolérance: question politique on morale?”, *Diogène* 176 (Paris, Editions Gallimard, 1996). Texto en inglés: “Toleration: A Political or Moral Question?”, en la correspondiente edición en inglés de *Diogenes* (Oxford, Berghahn Books) [IBD].
- “Dallo stato di natura alla genealogia”, *Studi Perugini* 2 (1996).
- “Shame, Guilt and the Structure of Punishment”, *Festschrift* para el Margrit Egner-Stiftung Prize [Premio Margrit Egner-Stiftung], Zurich, 1997.
- “Stoic Philosophy and the Emotions: Reply to Richard Sorabji”, en *Aristotle and After*, editado por R. Sorabji, Boletín del Instituto de Estudios Clásicos, Suplemento 68, Londres, Instituto de Estudios Clásicos, Escuela de Estudios Avanzados, Universidad de Londres, 1997.
- “Forward to Basics”, en *Equality*, editado por Jane Franklin (discusiones del reporte de la Comisión sobre Justicia Social, 1997), Londres, IPPR, 1997.
- “Plato against the Immoralist”, en *Platons Politeia*, editado por Otfried Höffe, Berlín, Akademie Verlag, 1997. [SP]
- “Moral Responsibility and Political Freedom”, *Cambridge Law Journal* 56 (1997). [PHD]
- “Berlin, Isaiah”, en *Routledge Encyclopedia of Philosophy*, editado por Edward Craig, Londres, Routledge, 1998.

“Virtues and Vices”, en *ibid.*

“Did Thucydides Invent Historical Time?”, *Jahrbuch 1996-97*, Wissenschaftskolleg zu Berlin, Berlín, Nicolaische Verlagsbuchhandlung, 1998, versión revisada, *Representations* 19 (2001).

Foreword to Angelika Krebs, *Ethics of Nature*, Berlín, Walter de Gruyter, 1999.

“Tolerating the Intolerable”, en *The Politics of Toleration in Modern Life*, editado por Susan Mendus, Edinburgo, Edinburgh University Press, 1999, Durham, Duke University Press, 2000. [PHD]

“In the Beginning Was the Deed”, en *Deliberative Democracy and Human Rights*, editado por Harold Hongju Koh y Ronald C. Slye, New Haven, Yale University Press, 1999. [IBD]

“Naturalism and Genealogy”, en *Morality, Reflection, and Ideology*, editado por E. Harcourt, Oxford, Oxford University Press, 2000.

“Philosophy as a Humanistic Discipline”, Third Annual Royal Institute of Philosophy Lecture, 2000, *Philosophy* 75 (2000). [PHD]

“Die Zukunft der Philosophie”, *Deutsche Zeitschrift für Philosophie* 48 (2000). Texto en inglés, “What Might Philosophy Become?” [PHD]

“Understanding Homer: Literature, History and Ideal Anthropology”, en *Being Human: Anthropological Universality in Transdisciplinary Perspectives*, editado por Neil Roughley, Berlín, Walter de Gruyter, 2000. [SP]

“Liberalism and Loss”, en *The Legacy of Isaiah Berlin*, editado por Mark Lilla, Ronald Dworkin y Robert Silvers, Nueva York, New York Review of Books, 2001.

“Some Further Notes on Internal and External Reasons”, en *Varieties of Practical Reasoning*, editado por Elijah Millgram, Cambridge, MIT Press, 2001.

“Foreword: Some Philosophical Recollections”, en *Wittgensteinian Themes: Essays in Honour of David Pears*, editado por David Charles y William Child, Oxford, Oxford University Press, 2001.

“From Freedom to Liberty: The Construction of a Political Value”, *Philosophy and Public Affairs* 30 (2001). [IBD]

Introducción a Friedrich Nietzsche, *The Gay Science*, editado por Bernard Williams, traducido por Josefine Nauckhoff. Cambridge, Cambridge University Press, 2001. [SP]

“Why Philosophy Needs History”, *London Review of Books*, octubre 17, 2002.

“Plato’s Construction of Intrinsic Goodness”, en *Perspectives on Greek Philosophy: S. V. Keeling Memorial Lectures in Ancient Philosophy 1991-2002*, editado por R. W. Sharples, Londres, Ashgate, 2003. [SP]

- “Relativism, History, and the Existence of Values”, en Joseph Raz *et al.*, *The Practice of Value*, editado por R. Jay Wallace, Oxford, Oxford University Press, 2003.
- “Unerträgliches Leiden”, en *Zum Glück*, editado por Susan Neiman y Matthias Kross, Berlín, Akademie Verlag, 2004. Texto en inglés: “Unbearable Suffering”. [SP]
- “Realism and Moralism in Political Theory”. [IBD]
- “The Liberalism of Fear”. [IBD]
- “Human Rights and Relativism”. [IBD]
- “Conflicts of Liberty and Equality”. [IBD]
- “Censorship”. [IBD]
- “Humanitarianism and the Right to Intervene”. [IBD]
- “Three Reasons for Talking about Collingwood”. [SP]
- “The Human Prejudice”. [PHD]

Reseñas seleccionadas

- Moral Judgement*, por D. Daiches Raphael, *Mind* (1957).
- The Contemplative Activity*, por Pepita Haezrahi, *Mind* (1957).
- Language and the Pursuit of Truth*, por John Wilson, *Times Literary Supplement*, 11 de enero de 1957 (publicado anónimamente).
- The Revolution in Philosophy*, por A. J. Ayer *et al.*, G. Ryle (ed.), *Philosophy* (1958).
- Aesthetics and Criticism*, por H. Osborne, *Mind* (1958).
- Plato Today*, por R. H. S. Crossman, *Spectator*, 11 de julio de 1959.
- English Philosophy since 1900*, por G. J. Warnock, *Philosophy* (1959).
- The Four Laves*, por C. S. Lewis, *Spectator*, 1 de abril de 1960.
- The Forms of Things Unknown*, por Herbert Read. *Spectator*, julio 29, 1960.
- Descartes, *Discourse on Method*, traducido por A. Wollaston. *Spectator*, 26 de agosto de 1960.
- The Liberal Hour*, por J. K. Galbraith; *Kennedy or Nixon?*, por Arthur Schlesinger, Jr. *Spectator*, 4 de noviembre de 1960.
- Thought and Action*, por Stuart Hampshire, *Encounter* (1960).
- Ethics since 1900*, por Mary Warnock, *Philosophical Books* (1960).

Sketch for a Theory of the Emotions, por J.P. Sartre, *Spectator*, 3 de agosto de 1962.

Sense and Sensibilia: Philosophical Essays, por J. L. Austin, *Oxford Magazine*, 6 de diciembre de 1962.

The Concept of a Person, por A. J. Ayer, *New Statesman*, 27 de septiembre de 1963.

Morals and Markets, por H. B. Acton, *Guardian*, 1 de abril de 1971.

Responsibility, por Jonathan Glover, *Mental Health* (1971).

Education and the Development of Reason, editado por R. F. Deardon, P. H. Hirst, y R. S. Peters, *Times Literary Supplement*, abril de 1972.

A Theory of Justice, por John Rawls, *Spectator*, 22 de junio de 1972.

Beyond Freedom and Justice, por B. F. Skinner, *Observer*, marzo de 1972.

Essays on Austin, editado por G. J. Warnock y J. O. Urmson, *New Statesman*, 31 de agosto de 1973.

“What Computers Cannot Do”, por Hubert Dreyfus, *New York Review*, octubre de 1973.

Essays on Wisdom, editado por Renford Bambrough, *Times Literary Supplement*, 31 de agosto de 1974.

The Socialist Idea, editado por Stuart Hampshire y L. Kolakowski, *Observer*, 5 de enero de 1975.

Anarchy, State, and Utopia, por R. Nozick, *Times Literary Supplement*, 17 de enero de 1975.

The Life of Bertrand Russell, por R. W. Clark; *The Tamarisk Tree: My Quest for Liberty and Love*, por Dora Russell; *My Father Bertrand Russell*, por Katharine Tait; y *Bertrand Russell*, por A. J. Ayer, *New York Review*, 4 de marzo de 1976.

Ethics of Fetal Research, por Paul Ramsey, *Times Literary Supplement*, 5 de septiembre de 1975.

The Selfish Gene, por R. Dawkins, *New Scientist*, 4 noviembre de 1976.

“Where Chomsky Stands”, *New York Review*, noviembre 11, 1976.

The Fire and the Sun, por Iris Murdoch, *New Statesman*, agosto 5, 1977.

Moore: G. E. Moore and the Cambridge Apostles, por Paul Levy, *Observer*, 28 de octubre de 1979.

Thinking, por G. Ryle, editado por K. Kolenda, *London Review of Books*, octubre de 1979.

Life Chances, por R. Dahrendorf, *Observer*, 27 de enero de 1980.

Rubbish Theory, por Michael Thompson, *London Review of Books*, febrero de 1980.

Lying, por Sissela Bok, *Political Quarterly* (1980).

Logic and Society and Ulysses and the Sirens, por Jon Elster, *London Review of Books*, 1 de mayo de 1980.

The Culture of Narcissism, por Christopher Lasch, *Nihilism and Culture*, por Johan Goudbloom, *London Review of Books*, 17 de julio de 1980.

Religion and Public Doctrine in England, por M. Cowling, *London Review of Books*, 2 de abril de 1981.

Nietzsche on Tragedy, por M. S. Silk, *Nietzsche: A Critical Life*, por Ronald Hayman; *Nietzsche, Volume 1: The Will to Power as Art*, por Martin Heidegger, traducido por David Farrell, *London Review of Books*, mayo de 1981.

After Virtue, por A. MacIntyre, *Sunday Times*, 22 de noviembre de 1981.

Philosophical Explanations, por R. Nozick, *New York Review of Books*, 18 de febrero de 1982.

The Miracle of Theism, por John Mackie, *Times Literary Supplement*, 11 de marzo de 1983.

Offensive Literature, por John Sutherland, *London Review of Books*, 17 de marzo de 1983.

Consequences of Pragmatism, por R. Rorty. *New York Review of Books*, 28 de abril de 1983, reimpresso en *Reading Rorty*, editado por Alan Malachowski, Oxford, Blackwell, 1990.

Collected Papers of Bertrand Russell, vol. 1, *Observer*, 22 de enero de 1984.

Reasons and Persons, por D. Parfit, *London Review of Books*, 7 de junio de 1984.

Critical Philosophy and Journal of Applied Philosophy. Times Higher Education Supplement, 15 de junio de 1984.

Wickedness, por Mary Midgley, *Observer*, 7 de octubre de 1984.

Secrets, por Sissela Bok; *The Secrets File*, por D. Wilson, *London Review of Books*, 18 de octubre de 1984.

Choice and Consequence, por Thomas Schelling, *Economics and Philosophy* (1985).

Privacy: Studies in Social and Cultural History, por Barrington Moore, Jr., *New York Review of Books*, 25 de abril de 1985.

Ordinary Vices, por Judith Shklar; *Immorality*, por Ronald Milo, *London Review of Books*, 6 de junio de 1985.

The Right to Know, por Clive Ponting; *The Price of Freedom*, por Judith Cook. *Times*

Literary Supplement, 4 de octubre de 1985.

Taking Sides, por Michael Harrington, *New York Times Book Review* [mediados de 1980].

A Matter of Principle, por Ronald Dworkin, *London Review of Books*, 17 de abril de 1986.

The View from Nowhere, por Thomas Nagel, *London Review of Books*, 7 de agosto de 1986.

The Society of Mind, por Marvin Minsky, *New York Review of Books*, 1987.

Whose Justice? Which Rationality?, por Alasdair MacIntyre, *Los Angeles Tribune*, 1988. Reimpreso: *London Review of Books*, 5 de enero de 1989.

Intellectuals, por Paul Johnson, *New York Review of Books*, 20 de julio de 1989.

Contingency, Irony, and Solidarity, por Richard Rorty, *London Review of Books*, noviembre de 1989.

Sources of the Self, por Charles Taylor, *New York Review of Books*, octubre de 1990.

Realism with a Human Face, por Hilary Putnam, *London Review of Books*, 7 de febrero de 1991.

The Saturated Self, por Kenneth J. Gergen, *New York Times*, 23 de junio de 1991.

Political Liberalism, por John Rawls, *London Review of Books*, 13 de mayo de 1993.

Inequality Reexamined, por Amartya Sen, *London Review of Books*, noviembre de 1993.

The Therapy of Desire, por Martha Nussbaum, *London Review of Books*, octubre de 1994.

Varios libros de Umberto Eco, *New York Review of Books*, 2 de febrero de 1995.

The Last Word, por Thomas Nagel, *New York Review of Books*, 19 de noviembre de 1998.

- [1] *Problems of the Self: Philosophical Papers 1956-1972*, Cambridge: Cambridge University Press, 1973. Hay traducción al castellano: *Problemas del yo*, México, UNAM, 1986.
- [2] *Moral Luck: Philosophical Papers 1973-1980*, Cambridge, Cambridge University Press, 1981.
- [3] *Making Sense of Humanity and Other Philosophical Papers 1982-1993*, Cambridge, Cambridge University Press, 1995.
- [4] Myles Burnyeat (coord.), *The Sense of the Past: Essays in the History of Philosophy*, Princeton, Princeton University Press, 2006.
- [5] Geoffrey Hawthorn (coord.), *In the Beginning was the Deed: Realism and Moralism in Political Argument*, Princeton, Princeton University Press, 2005.
- [6] Para una discusión sobre esta idea de importancia, véase *Ethics and the Limits of Philosophy*, Londres: Fontana, 1985, pp. 182ss.
- [7] Ensayo XVI, p. 205.
- [8] *Ethics and the Limits of Philosophy*, p. 139.
- [9] Esta es una cita de John McDowell, "Critical Notice of *Ethics and the Limits of Philosophy*", *Mind* 95, 1986, p. 380.
- [10] Cfr. p. 66.
- [11] Cfr. p. 214.
- [12] Véase, e.g., *Ethics and the Limits of Philosophy*, p. 140.
- [13] Véase, e.g., *ibid.*, p. 135.
- [14] Creo que la relación entre esto y el título de la pintura de Gauguin usada como portada para *Moral Luck*, a saber *D'où venons nous... Que sommes nous... Où allons nous*, no es un accidente.
- [15] *Ethics and the Limits of Philosophy*, p. 197.
- [16] Ensayo XVI, p. 218.
- [17] *Ethics and the Limits of Philosophy*, pp. 168-170.
- [18] *Ibid.*, pp. 170-171.
- [19] Platón, *Apología de Sócrates*, 38a.
- [20] *La ética y los límites de la filosofía*, p. 170.
- [21] Estoy muy agradecido con Miranda Fricker, Geoffrey Hawthorn, Ian Malcolm, y especialmente con Patricia Williams por sus consejos y aliento.
- [22] *Truth and Truthfulness; an Essay in Genealogy*, Princeton, Princeton University Press, 2002.
- [23] Son publicados aquí exactamente como aparecieron en su versión original, excepto

por alguna estandarización mínima y la corrección de errores menores (los cuales he señalado siempre que la corrección me pareció algo más que de rutina). Estoy muy agradecido con Lauren Lepow, directora de edición en la Princeton University Press, por su ayuda en la identificación de estos errores.

- [24] Hay una traducción al alemán de este ensayo, publicado bajo el título “Die Zukunft der Philosophie”, en *Deutsche Zeitschrift für Philosophie* 48, 2000.
- [a] Este artículo, sustancialmente en su forma actual, fue leído en mayo de 1954 para el Club Socrático de la Universidad de Oxford; debo expresar mi gratitud al presidente del club y al editor de su publicación, el *Socratic Digest*, por permitir que se publicara aquí.
- [b] Ésta es una referencia a los editores Anthony Flew y Alasdair MacIntyre por *New Essays in Philosophical Theology* (Nuevos ensayos en la teología filosófica, Londres, SCM Press, 1955), en el cual apareció por primera vez el ensayo de Williams. [Coord.]
- [25] Para un argumento similar véase el ensayo XI más adelante. Esta es una referencia a C. B. Martin, “The Perfect Good”. [Coord.]
- [26] La conexión de esto con la idea de un alma inmortal es obvia, y es importante, pero no puede seguirse aquí.
- [27] Pero desde luego no podemos entender correctamente la primera parte de la paradoja si no entendemos la segunda.
- [28] Esta frase, y la idea subyacente, vienen de Wittgenstein. El exponente más completo de esa teoría es John Wisdom. Véase *Philosophy and Psychoanalysis*, Blackwell, Oxford, 1953.
- [29] *Eth. Nic.* 1174^a 13 ss.
- [30] “Pleasure”, *P:A:S: Supp.*, vol. XXVIII (1954), pp. 135ss. No estoy seguro de si Ryle contribuyó a esta perspectiva. Algunos comentarios en este artículo sugieren que no, pero la concentración en el caso de la actividad, en gran parte de sus argumentos, está en contra de aceptar el placer como un sentimiento.
- [31] Aunque no deberíamos desestimar nuestras capacidades en esta dirección; cf. Price “Belief and Will”, *P:A:S: Supp.*, vol XXVIII (1954), Inaugural Address.
- [32] El punto es señalado varias veces por Wittgenstein en *Philosophical Investigations*.
- [33] Esto *puede* estar relacionado con un punto señalado por Miss G.E.M. Anscombe acerca de ciertos movimientos corporales: cf. *Intention* (Blackwell, 1957) para. 8. Su frase “cause known without observation” (“causa conocida sin observación”) no es adecuada para el presente caso; pero no estoy seguro tampoco de que sea una descripción muy satisfactoria del caso que ella está presentando.
- [34] Cf. Platón, *Filebo* 39d ss.
- [35] *Filebo* 50b ss.
- [36] *Ibid.*, pp. 139ss.; para el placer como un modo de atención, véase la página 142.
- [37] Esta dificultad ha sido bien discutida por Gettier. Véase su artículo reimpreso en Phillips Griffiths (coord.), *Knowledge and Belief*, Oxford, 1967, y también la introducción de Griffiths en ese mismo libro. La presente discusión también está en

- deuda con Martin y Deutscher, “Remembering”, *Philosophical Review*, 1966.
- [38] Nigel Dennis, *Cards of Identity*, Weidenfeld and Nicolson, Londres, 1955, pp. 40-41.
- [39] *De Corpore* 2.11. Estoy en deuda sobre esta pregunta con David Wiggins: véase Wiggins, 1980, p. 92. [Esta es una referencia a David Wiggins, *Sameness and Substance*, Blackwell, Oxford. E.]
- [40] Es defendida (bajo el nombre de la “teoría del continuante más cercano”) por Robert Nozick (1981, capítulo 1). [Esto es una referencia a Robert Nozick, *Philosophical Explanations*, Harvard University Press, Cambridge. E.]
- [41] Citado en Abrams y Hogg 1990, p. 2. [Esta es una referencia a Dominic Abrams y Michael A. Hog (eds.), *Social Identity Theory*, Hemel Hempstead, Harvester. E.]
- [c] MENSA es un sociedad británica de personas que piensan que tienen un alto coeficiente intelectual. MENSA proporciona acertijos y desafíos intelectuales a sus miembros, así como la posibilidad de convivir con personas afines. Véase <http://www.mensa.org.uk/>. [T.]
- [42] *Principia Ethica*, Cambridge, 1903, secciones 89 y 117. Moore posteriormente se retractó de su opinión.
- [43] He argumentado en favor de esto en “Formal and substantial individualism”, *Proceedings of the Aristotelian Society*, LXXXIV, 1984-1985.
- [44] El punto en que es necesario que tales teorías no sean inconsistentes está bien planteado por Derek Parfit (*Reasons and Persons*, Oxford University Press, 1984), quien distingue varios modos en que las teorías pueden excluir la creencia, o una creencia ampliamente difundida y sincera en sí misma. Sin embargo, esto tiene menos interés para las implicaciones sociales de estos resultados, una cuestión que me parece fundamental. He discutido el ideal de transparencia y asuntos relacionados en *Ethics and the Limits of Philosophy* (Londres, Collins, Cambridge, Mass., Harvard University Press, 1985), especialmente en pp. 101 y ss.
- [45] Véase David Pears, “Courage as a mean”, en Amelie Rorty (ed.), *Essays on Aristotle’s Ethics*, Berkeley: University of California Press, 1980.
- [46] D. Hume, *A Treatise of Human Nature*, L. A. Selby-Bigge (ed.) (Oxford: Oxford University Press, 1975), libro III, parte II, sección 12.
- [47] Cf. *Ethics and the Limits of Philosophy*, *op. cit.*, sobre el tema del conocimiento a través de conceptos “densos”. He sugerido allí que la posibilidad de tal conocimiento tiende a retroceder ante la crítica reflexiva; pero este es, por supuesto, un punto separado.
- [48] Quizá valga la pena agregar que solamente si el objetivismo se entiende, estrechamente, en términos del realismo, es que una concepción de la disposición tiene que ser no objetiva. Tal concepción puede buscar muy consistentemente una fundamentación objetiva de la vida ética tratando de mostrar que una vida caracterizada por alguna estructura de disposiciones éticas es la vida más digna de ser vivida.
- [49] Para una discusión ulterior de este punto, véase “Formal and substantial

individualism”, *op. cit.*

- [*] Aquí y en lo que sigue, “Williams, 1985” es una referencia a Bernard Williams, *Ethics and the Limits of Philosophy*, Fontana, Londres. [E.]
- [**] Esto es una referencia a Douglas Seanor y N. Fotion (eds.), *Hare and Critics: Essays on Moral Thinking*, Oxford University Press, Oxford, 1988, en el que el ensayo de Williams apareció por primera vez y en el cual hay una réplica de Hare. [E.]
- [50] Discutiré la teoría presentada en *MT*. [De aquí en adelante *MT* es una referencia a R. M. Hare, *Moral Thinking: Its Levels, Methods and Point*, Oxford University Press, Oxford, 1981. E.] No me ocuparé del desarrollo de los puntos de vista de Hare, excepto en 7.
- [51] Hare hace suposiciones simplificadoras acerca de la clase de preferencias a ser considerada. Véase *MT*, capítulo 5. Este es un asunto importante, que plantea preguntas acerca de la motivación de las teorías utilitaristas en general (*cf.* Williams, 1985, pp. 86-89; Sen y Williams, 1982, introducción), pero no las proseguiré aquí. [Lo segundo es una referencia a Amartya Sen y Bernard Williams (eds.), *Utilitarianism and Beyond*, Cambridge University Press Cambridge. E.]
- [52] Para esta objeción, véase Mackie, 1977, p. 97; Nagel, 1982. [Estas son referencias respectivamente a J. L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin), y Thomas Nagel, “Review of *Moral Thinking*, por R. M. Hare”, *London Review of Books*, 15 de enero. E.] Sobre el exigente concepto de Hare acerca de la preocupación imparcial por las preferencias de los demás, véase Williams, 1985, pp. 89-92.
- [53] Si una teoría utilitaria no ha de plantear la dificultad en absoluto, no debería ir mucho más allá del “modelo del recibo del gas” discutido en Williams, 1972. [Esta es una referencia a Bernard Williams, *Morality: An Introduction to Ethics*, Cambridge University Press, Cambridge. E.]
- [54] Parfit, 1984, parte I. [Esta es una referencia a Derek Parfit, *Reasons and Persons*, Oxford University Press, Oxford. E.]
- [55] El punto es explícito en *MT* 17ss. Para el muy especial carácter de la investigación lingüística, *cf.* Williams, 1985, capítulo 7.
- [d] Esta es una referencia a A. J. Ayer, *Russell*, Fontana, Londres, 1972. [E.]
- [56] “On the Analysis of Moral Judgements”, publicado originalmente en *Horizon*, vol. XX (1949), reimpresso en *PE*. [Esto es una referencia a A. J. Ayer, *Philosophical Essays*, Macmillan, Londres, 1954. E.]; esta cita es del reimpreso, pp. 237-238.
- [57] “Philosophy and Politics”, publicada en 1967 por la Liverpool University Press, reimpresso en *MCS*. [Esta es una referencia a A. J. Ayer, *Metaphysics and Common Sense*, Macmillan, Londres, 1969. E.]; la cita es del reimpreso, pp. 259-260.
- [e] Esta es una referencia a Bernard Williams, *Ethics and the Limits of Philosophy (Ética y límites de la filosofía)*, Fontana, Londres. [E.]
- [58] Michael S. Moore, *Act and Crime: The Philosophy of Action and its Implications for Criminal Law*, 1993. [El ensayo de Williams apareció por primera vez en un

número de *University of Pennsylvania Law Review* que se componía de ensayos encargados para un simposio sobre el libro de Moore. Éste está publicado por Oxford University Press. E.]

[59] Véase *ibid.* pp. 4-5.

[60] Véase Bernard Williams, *Shame and Necessity*, pp. 55-56 (1993).

[61] Moore, *op. cit.*, p. 263.

[62] Véase *id.*, pp. 280-292.

[63] Véase Aristóteles, *Ethica nichomachea* 1094 b 12.

[64] Véase H.L.A. Hart, *Punishment and Responsibility: Essays in the Philosophy of Law*, p. 109 (1968).

[65] Véase Bernard Williams, *Voluntary Acts and Responsible Agents*, 10 *Oxford J. Legal Stud.* 1, 1 (1990).

[66] Moore, *supra* nota 1, en 253.

[67] Véase Williams, *supra* nota 8, en 1.

[68] William Shakespeare, *Macbeth* acto 5, escena 1.

[69] Moore, *supra* nota 1, p. 254.

[70] *Idem.*

[71] *Idem.*

[72] Véase P.F. Strawson, *Individuals: An Essay in Descriptive Metaphysics*, 1959, pp. 87-116.

[73] He argumentado esto anteriormente. Véase Bernard Williams, *Are Persons Bodies?*, en *Problems of the Self: Philosophical Papers, 1956-1972*, 1973, pp. 64-70.

[f] He reconstruido ligeramente la cláusula final de esta proposición, la cual fue alterada en el original. [E.]

[74] Moore, *supra* nota 1, en 255.

[75] *Idem.*

[76] Véase Donald Davidson, *Agency*, en *Essays on Actions and Events* 43-61 (1980). La perspectiva de Davidson es que un evento constituye una acción sólo si hay alguna descripción bajo la cual esa acción sea intencional.

[77] No me ocuparé aquí de este argumento. Lo que es importante es que es más fuerte que una mera aplicación del argumento de la ilusión. Cuando el agente hace lo correcto, está en el mismo estado cognitivo inusual (sujeto a las sugerencias del hipnotizador) que cuando hace algo mal. A pesar de todo, el argumento aún depende de suposiciones disputables en la teoría del conocimiento.

[78] Esto no es negar que hay una cuestión importante acerca de qué está mal exactamente con tales agentes; la naturaleza, por ejemplo, de su disociación. Tales cuestiones son discutidas por Stephen Morse en su contribución a este volumen. Véase en forma general Stephen J. Morse, *Culpability and Control*, 142 *U. Pa. L. Rev.* 1587 (1994). [Esta es una referencia a la *University of Pennsylvania Law Review*. —E.]

[79] Seguiré la historia que se presenta en casi toda la película. Resulta que al final de la película esta historia es realmente una fantasía del joven psicótico que ha estado

narrándola.

- [80] Moore da este tipo de interpretación. Véase Moore, *supra* nota 1, en 236. La apelación a lo que él llama “matices individuales”, sin embargo, seguramente no la hará para “apuñalar” y “asfixiar”.
- [81] Pero no lo hizo, en términos de la película, para su propósito avasallador: él estaba experimentando con el sonambulismo.
- [82] Véase *supra* nota 3 y el texto que la acompaña.
- [83] Moore, *supra* nota 1, en 226-232.
- [g] Dado que la descripción de este caso que hace Williams omite, de manera un tanto desconcertante, cualquier razón que Maurice pudiera tener para preferir ir a las montañas a quedarse en casa, y dado que lo que Broome dice acerca de esto desmiente la caracterización que hace Williams de Maurice como “anticultura”, pienso que sería útil citar, de manera completa, la propia descripción que hace Broome del caso (1991*b*, p. 101): “El montañismo asusta [a Maurice], así que prefiere visitar Roma. Andar de turista mirando lo aburre, así que prefiere quedarse en casa. Pero quedarse en casa cuando pudiera haber ido a hacer montañismo sería, él cree, cobarde. Ésta es la razón por la que, si tuviera que hacer la elección entre quedarse en casa e ir de montañista, escogería ir de montañista. (Visitar Roma cuando pudiera haber ido de montañista le parece a él, más bien, cultural más que cobarde.)” [E.]
- [84] Susan Hurley (1989), quizá de manera engañosa, describe cualquier posición que reconoce esto como “objetivista”.
- [h] He agregado tanto la palabra “abandonamos” como la frase “la idea de que hay” dentro del texto original porque me parece que le da un mejor sentido al texto. El ensayo original contenía un gran número de errores tipográficos, y creo que es más que posible que algunas de esas expresiones se hayan quitado del propio manuscrito de Williams. [E.]
- [i] La conferencia Aquinas de la Sociedad Internacional de Criminología, dada en Cambridge el 5 de mayo de 1994, tuvo lugar en el Instituto de Criminología de la Universidad de Cambridge.
- [85] Williams, *Shame and Necessity*, Berkeley 1993, capítulo 3.
- [86] Se ilustra en *Shame and Necessity* mediante el contraste entre dos agentes en Homero: Telémaco, quien descuidadamente deja una puerta abierta, y Agamenón, quien aleja a Briseida de Aquiles y, desde luego, quería hacerlo, pero se encontraba (él después lo afirma) en un estado mental extraño cuando hizo eso.
- [87] Moore, *Act and Crime: The Philosophy of Action and Its Implications for Criminal Law*, Oxford, 1993, pp. 253 ss. Para una discusión de esto y concepciones afines de Moore véase el simposio en *University of Pennsylvania Law Review* 142 (1994), pp. 1443-1840. [Ésta es la fuente del ensayo 9 antes mencionado. E.]
- [88] La versión voluntarista. La retribución no está confinada a lo voluntario: véase p. 145.
- [89] Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (Oxford

- 1968), especialmente el capítulo 2.
- [90] Estoy agradecido con Antony Duff por tocar este punto en su exposición en la conferencia Aquinas.
- [91] “Lo hizo cambiar de opinión con respecto a los valores” es la frase de Robert Nozick en Nozick, *Philosophical Explanations* (Harvard 1984), en pp. 363ss. Para ulteriores comentarios sobre las presentes líneas, véase “Nietzsche’s minimalist moral psychology”, *European Journal of Philosophy* 1.1 (1993), reimpresso en Williams, *Making Sense of Humanity* (Cambridge 1995). [Esto está ahora también reimpresso en Bernard Williams, *The Sense of the Past: Essays in the History of Philosophy*, ed. Myles Burnyeat, Princeton University Press, 2006. E.]
- [92] Nozick, *Anarchy, State and Utopia*. Basic Books 1974, p. 225.
- [93] Véase John Rawls, *A Theory of Justice*, Oxford, Oxford University Press, 1971; John Rawls, *Political Liberalism*, Nueva York, Columbia University Press, 1993; Thomas Nagel, *Equality and Impartiality*, Oxford, Oxford University Press, 1991; Ronald Dworkin, “What is Equality?”, *Philosophy and Public Affairs*, 10, 1981.
- [94] Nagel, *Equality and Impartiality*, p. 156.
- [95] Judith Shklar, “The Liberalism of Fear” en Nancy Rosenblum (ed.), *Liberalism and Moral Life*, Cambridge, Mass, Harvard University Press, 1989, pp. 21-38.
- [96] Agradezco aquí la ayuda de Jill Kraye, “Moral Philosophy”, en C.B. Schmitt (ed.), *The Cambridge History of Renaissance Philosophy*, Cambridge, Cambridge University Press, 1988, esp., pp. 306-316.
- [97] Friedrich Nietzsche, “On Truth and Lies in a Nonmoral Sense”, en Daniel Breazeale (trad. y ed.), *Philosophy and Truth Solutions from Nietzsche’s Notebooks of the Early 1870s*, Harvester Press, Brighton, 1979, primer párrafo.
- [98] Para una versión teóricamente poco ambiciosa de un ensayo de “teoría crítica” que se aplica a tales situaciones, véase mi *Truth and Truthfulness: An Essay in Genealogy*, Princeton University Press, Princeton, 2002, capítulo 9, secciones 4 y 5.
- [99] Cf. en esta conexión la opinión del ya fallecido Stephen J. Gould acerca de la falsa impresión de “progreso” dada por la vieja representación estándar del árbol evolutivo.
- [100] Peter Singer, *Unsanctifying Human Life: Essays on Ethics [UHL]*, ed. Helga Kuhse, Oxford, Blackwell, 2002, p. 193. [Esta cita es de un artículo titulado “Individuals, Humans, and Persons: The Issue of Moral Estatus”, coautoría con Helga Kuhse. E.]
- [101] *UHL*, p. 239. [Esta cita es de un artículo titulado “Should All Seriously Disabled Infants Live?”, coautoría con Helga Kuhse. E.]
- [102] *UHL*, p. 194. [Véase antes, n. 5. E.] Relativo a potencialidad, véase Peter Singer y Karen Dawson, “Technology and the Argument from Potential”, en *UHL*, pp. 199-214.
- [103] *UHL*, p. 192. [Véase antes, n. 5. E.]
- [104] Ronald Dworkin, en *Life’s Dominion: Argument about Abortion and Euthanasia* (Londres: Harper Collins, 1993), trata de reclutar “la vida es sagrada” en favor de políticas radicales. Dudo que esto funcione algo mejor.

- [105] Roderick Firth, "Ethical Absolutism and the Ideal Observer", *Philosophy and Phenomenological Research* 12 (1952): 317-45.
- [106] *Ethics and the Limits of Philosophy* (Londres: Fontana, 1985), pp. 118-119.
- [107] Robert Nozick, *Anarchy, State, and Utopia*, Nueva York, Basic Books, 1974, pp. 45-47. [Williams originalmente tomó como referencia *Philosophical Explanations* de Nozick, pero pienso que este pasaje de *Anarchy, State, and Utopia* debe haber sido lo que él tenía en mente. E.].
- [108] *Der Einziger und sein Eigenthum*, traducido por S. T. Byington como *The Ego and His Own*, ed. James J. Martin (Sun City, Calif.: West World Press, 1982), p. 128.
- [109] En esta perspectiva, podríamos considerar el hecho de que Kant, pese a su énfasis en la aplicación de la ley moral a los agentes racionales como tales, expresa la tercera formulación del imperativo categórico en razón de cómo debemos asumir la noción de *humanidad*.
- [110] Esto es implicar que el estudio de las interrelaciones conceptuales de la perspectiva de un grupo forma parte del estudio científico social de ese grupo. No implica que forme más que una parte. La idea de que la ciencia social es, más que todo lo demás, investigación conceptual (*cf.* Winch), es una posición muy diferente y para mí, inaceptable.
- [111] Debo este argumento a T. Williamson, quien lo desarrolla en una forma más precisa y rica en su libro *Vagueness*, Londres, 1995.
- [j] Dictada como la conferencia anual del Royal Institute of Philosophy el 23 de febrero de 2000.
- [112] El punto, en particular, de hacer que lo familiar parezca extraño, y a la inversa. He dicho algo más acerca de esto en "Descartes and the Historiography of Philosophy", en John Cottingham (ed.), *Reason, Will and Sensation*, Oxford, Clarendon Press, 1994. [Esto ha sido reimpresso en Bernard Williams, *The Sense of the Past: Essays in the History of Philosophy*, ed. Myles Burnyeat, Princeton, Princeton University Press, 2006. E.]. La referencia a Collingwood es a *An Autobiography*, Oxford, Clarendon Press, 1939, p. 63 ss.
- [113] La primera concepción fue expresada, en una forma vulgarizada, en la literatura del "positivismo terapéutico". La segunda se halla ricamente desarrollada en la obra de Stanley Cavell.
- [114] Cambridge MA, Harvard University Press, 1992, prefacio, p. v.
- [115] *Ibid.*, p. 108.
- [116] *Descartes: The Project of Pure Enquiry* (Harmondsworth: Penguin, 1978), p. 64.
- [117] *Renewing Philosophy*, p. 123.
- [118] Una sobresaliente discusión se halla en A. W. Moore, *Points of View*, Oxford, Clarendon Press, 1997.
- [119] Éste es el punto que debería ser relevante a la cuestión de si la filosofía formaría parte del contenido del saber absoluto. Más aún, si Putnam quiso decir que cualquier afirmación, por el hecho de contener términos regulados por relaciones semánticas normativas es normativa, tendría que decir que todas las afirmaciones son

normativas.

[120] *Descartes*, pp. 300-303.

[121] Una línea, más bien, similar fue adoptada por algunos defensores de la religión a principios de la revolución científica.

[122] Andrew Pickering, *La construcción de los quarks* (Edinburgh University Press, 1984). Debería decirse que la historia de Pickering plantea preguntas importantes acerca de la interpretación del “descubrimiento” de los quarks.

[123] *Contingency, Irony and Solidarity*, Cambridge University Press, 1989, especialmente los capítulos 3 y 4.

[124] Ésta es (en efecto) una aseveración central del libro de Thomas Nagel *The Last Word* (Oxford University Press, 1997). Sus argumentos inciden estrechamente en la presente discusión. He comentado sobre ellos en una reseña del libro, *New York Review of Books* XLV, 18 (19 de noviembre de 1998).

[125] Véase e. g. “Wittgenstein and Idealism”, reimpresso en *Moral Luck*, Cambridge University Press, 1981. [También está reimpresso en *The Sense of the Past*. E.]. La cuestión del idealismo no es relevante en el contexto actual.

[126] Esto necesita cualificación con respecto al trabajo más reciente de Rawls, quien exhibe un sentido más fuerte de contingencia histórica que el que se hallaba presente en *A Theory of Justice*.

[127] *Daybreak*, sec. 542.

[128] Este ensayo está basado en una conferencia inaugural dada en el Centro para la Filosofía Postanalítica, Universidad de Southampton, en noviembre de 1997.

[129] Oxford: Clarendon Press, 1990.

[130] David Wood, ahora en la Universidad Vanderbilt, ha acuñado el útil eslogan “Postanalítico = Post continental”.

[131] *Origins of Analytical Philosophy*, Londres, Duckworth, 1993. Dummett mismo, aunque su libro pretende solamente discutir “orígenes” de la filosofía analítica, no “los orígenes”, tiende mucho a identificar el estilo analítico con el “sesgo lingüístico”, el método de tratar el lenguaje como explicativamente anterior al pensamiento. Es seguramente un problema para esta explicación que la cuestión de la prioridad entre lenguaje y pensamiento ha sido en sí misma un asunto central en la filosofía analítica.

[132] *Limited Inc*, ed. G. Graff, Evanston: Northwestern University Press, 1988. [Es probable que Williams específicamente tuviera en mente “Limited inc a b c ...”, reimpresso ahí. (La traducción alemana de este ensayo no incluye esta nota, así que no hay corroboración ahí). E.].

[133] Véase en particular *Contingency, Irony and Solidarity* (Cambridge: Cambridge University Press, 1989).

[134] “Desde luego”: pero *cf.* el excelente artículo de Lorraine Daston. [Daston es un escritor prolífico y no hay manera de estar seguro a cuál artículo se estaba refiriendo Williams. (La traducción alemana de este ensayo no incluye esta nota, así que no provee ninguna ayuda con este problema). E.].

[135] *The Gay Science*, ed. Bernard Williams, trad. Josefine Nauckhoff, Cambridge:

Cambridge University Press, 2001, p. 344. No estoy en lo más mínimo sugiriendo (ni lo hizo Nietzsche) que la idea de confianza plena, en la ciencia o en la filosofía, sea irreal. El punto es que lo ideal, lo cual juega una parte importante en las controversias acerca de lo analítico como opuesto a otros estilos en filosofía, sorprendentemente no es examinado.

[136] Véase en particular *The Claim of Reason*, Oxford, Clarendon Press, 1979, e *In Quest of the Ordinary*, Chicago, University of Chicago Press, 1988.

[137] *Philosophical Investigations*, trad. G. E. M. Anscombe, ed. rev. (Oxford: Blackwell, 1974), énfasis en el original.

[138] Cf. *Feierstag*, una festividad pública. Como yo lo entiendo, *feiern* no sólo se refiere estandarizadamente a irse de vacaciones, o tomarse un descanso; la idea central es quizá que las reglas o expectativas usuales se están relajando.

[139] Bajo una amplia concepción del significado, estrechamente relacionada con el uso. Es consistente con esta aproximación a la filosofía aceptar (aunque los seguidores de Wittgenstein tienden a no hacerlo) que hay aplicaciones más estrechas de “carente de significado”, relacionadas con la incoherencia sintáctica o semántica, en las que el lenguaje en cuestión aquí no es carente de significado.

[140] *Zur Genealogie der Moral*, 2.13: “definierbar ist nur das, was keine Geschichte hat”. (El contexto es una discusión del castigo.)

[141] *Filósofos mujeres*, (Londres, Dent, 1996, pp. xxxiii-xxxiv. También sostiene que negar esto (o al menos negarlo en ciertos estilos) no es hacer filosofía: “Aquellos que [...] argumentan, como los posmodernos, [...] que no hay mundo común compartido, sino que cada uno de nosotros construimos nuestro propio mundo [...] están relacionados [...] no con la filosofía, sino con una especie de antropología”. Esto plantea el mismo tipo de problema que la posición mencionada en la nota 4 anterior: ¿cómo identificamos una cuestión filosófica? Difícilmente puede ser que una cuestión sea filosófica si obtiene una respuesta afirmativa, y no si obtiene una respuesta negativa. (Es irónico, en relación con la posición de Mary Warnock, que un filósofo, quizá el único, en haber explícitamente considerado esta dificultad es Collinwood, con su plenamente historicista teoría de las presuposiciones absolutas que, precisamente, no son respuestas a preguntas).



¿Qué puede –y qué no puede– hacer la filosofía? ¿Cuáles son sus riesgos éticos y sus posibles recompensas? ¿En qué difiere de la ciencia? En *La filosofía como una disciplina humanística*, Bernard Williams aborda estas preguntas y presenta una interesante visión de la filosofía como fundamentalmente diferente de la ciencia en sus metas y métodos. Escrito con su distintiva combinación de rigor, imaginación, profundidad y humanismo, el libro muestra por qué Williams fue uno de los filósofos más grandes del siglo xx.

Desde su primera publicación hasta una de sus últimas conferencias, los ensayos de este libro, no publicados ni compilados previamente, abordan la metafísica, la epistemología y la ética, así como el alcance y los límites de la filosofía misma. Los ensayos forman una unidad por la constante preocupación de Williams por que la filosofía mantenga contacto con los problemas humanos. En la introducción, A. W. Moore escribe que el ensayo que presta título al libro es “una especie de manifiesto de la concepción de Williams sobre el trabajo de su propia vida”. Es donde más directamente aborda la cuestión de “cómo la filosofía puede o no puede contribuir al proyecto de encontrarle sentido a las cosas”; respondiendo que lo que la filosofía puede encontrar es el sentido del “ser humano”.

La filosofía como una disciplina humanística es uno de los tres libros póstumos de Williams que el FCE publica del filósofo. En el principio era el hecho: realismo y moralismo en el argumento político y El sentido del pasado: ensayos de historia de la filosofía son los títulos de los otros dos.

Índice

Índice	5
Prefacio	6
Introducción	9
Primera Parte. Metafísica y epistemología	17
I. La paradoja de Tertuliano	18
II. Argumentos metafísicos	36
III. Placer y creencia	47
IV. Conocimiento y razones	58
V. Identidad e identidades	67
Segunda Parte. Ética	74
VI. La primacía de las disposiciones	75
VII. La estructura de la teoría de Hare	83
VIII. Subjetivismo y tolerancia	92
IX. El actus reus del doctor Caligari	102
X. Los valores, las razones y la teoría de la persuasión	112
XI. Responsabilidad moral y libertad política	121
XII. Tolerar lo intolerable	127
XIII. El prejuicio humano	135
Tercera Parte. El alcance y los límites de la filosofía	151
XIV. La filosofía política y la tradición analítica	152
XV. La filosofía y el entendimiento de la ignorancia	165
XVI. La filosofía como una disciplina humanística	175
XVII. ¿En qué podría convertirse la filosofía?	192
Bernard Williams Publicaciones filosóficas completas	204
Notas	220