

**DONALD DAVIDSON**

**ENSAYOS  
SOBRE ACCIONES  
Y SUCESOS**



Instituto de Investigaciones Filosóficas  
UNAM  
CRÍTICA

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del *copyright*, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo públicos.

Título original:  
ESSAYS ON ACTIONS AND EVENTS  
Oxford University Press, Oxford

Cubierta: Enric Satué, sobre un trabajo artesanal, en pan, de Eduardo Crespo

© 1980: Donald Davidson

© 1995 de la traducción castellana: Instituto de Investigaciones Filosóficas de la Universidad Nacional Autónoma de México

© 1995 de esta edición para México: Instituto de Investigaciones Filosóficas de la Universidad Nacional Autónoma de México, Circuito Mario de la Cueva, Ciudad de la Investigación en Humanidades, Coyoacán, 04510 México, D.F.

© 1995, de esta edición para España y América, excepto México:

CRÍTICA (Grijalbo Mondadori, S. A.), Aragón, 385, 08013 Barcelona

ISBN: 84-7423-660-6

Depósito legal: B. 30.329-1995

Impreso en España

1995. – NOVAGRÁFIK, Puigcerdà, 127, 08019 Barcelona

## ORIGEN DE LOS ENSAYOS Y RECONOCIMIENTOS

El ensayo 1, «Acciones, razones y causas», se presentó en un simposio sobre acción en la reunión de la American Philosophical Association en 1963 y se publicó en el *Journal of Philosophy* 60, 1963. Se reimprime aquí con permiso de los editores.

El ensayo 2, «¿Cómo es posible la debilidad de la voluntad?», se publicó primero en *Moral Concepts*, editado por Joel Feinberg, Oxford Readings in Philosophy, 1970. Se leyeron versiones anteriores del artículo en el Annual Oregon Colloquium in Philosophy de 1967, ocasión en la que me ayudaron los comentarios cuidadosos de Donald G. Brown, y en el Chapel Hill Colloquium in Philosophy, en el que escuché las hábiles críticas de Gilbert Harman.

El ensayo 3, «De la acción» [«Agency»], se presentó en el cuarto coloquio de filosofía que tuvo lugar en la Universidad de Western Ontario en noviembre de 1968, y se publicó en Robert Binkley, Richard Bronaugh, Ausonio Marras (comps.), *Agent, Action, and Reason*, University of Toronto Press, 1971. James Cornman comentó el artículo en el coloquio y la versión impresa mejoró con su consejo. Se reimprime aquí con la autorización de la University of Toronto Press.

El cuarto ensayo, «Libertad para actuar», fue publicado primero en *Essays on Freedom of Action*, compilado por Ted Honderich, Routledge and Kegan Paul, 1973 (se reimprimió en edición de bolsillo en 1978).

El ensayo 5, «Tener la intención», se leyó en el primer Jerusalem Philosophical Encounter en diciembre de 1974, y Stuart Hampshire

me hizo una réplica animada. Mi artículo y su réplica se publicaron en Yirmiahu Yovel (comp.), *Philosophy of History and Action*, editado por D. Reidel y The Magnes Press, The Hebrew University, 1978. Una versión anterior se presentó en el Chapel Hill Colloquium in Philosophy en octubre de 1974, en donde recibió un examen riguroso por parte de Paul Grice.

«La forma lógica de las oraciones de acción», el ensayo 6, se presentó en un congreso sobre la lógica de la decisión y de la acción en la Universidad de Pittsburgh en marzo de 1966. E.J. Lemmon, H.N. Castañeda y R.M. Chisholm hicieron comentarios, y mi artículo, sus comentarios y mis respuestas se publicaron en Nicholas Rescher (comp.), *The Logic of Decision and Action*, University of Pittsburgh Press, 1967. Los orígenes del material adicional al final del ensayo 6 se especifican en éste.

El ensayo 7, «Relaciones causales», se presentó en un simposio del mismo nombre en la reunión de la American Philosophical Association Eastern Division en diciembre de 1967; Zeno Vendler hizo la réplica. Nuestros dos artículos se publicaron en el *Journal of Philosophy* 64, 1967, y mi ensayo se reimprime con la autorización de los editores.

El siguiente ensayo, el número 8, «La individuación de los sucesos», se publicó en Nicholas Rescher (comp.), *Essays in Honor of Carl Hempel*, D. Reidel, 1969, pp. 216–234, copyright ©1969 por D. Reidel Publishing Company, Dordrecht, Holanda. Se reimprime con la autorización de D. Reidel Publishing Company. David Kaplan comentó una versión anterior que leí en un coloquio en la Universidad de California en Irvine en abril de 1967. Parte de su buen juicio se ha incorporado a la versión impresa.

Los ensayos 9 y 10, «Los sucesos como particulares» y «Sucesos eternos vs. sucesos efímeros», fueron escritos con motivo de un simposio sobre sucesos en la reunión de mayo de 1970 de la Western Division de la American Philosophical Association. Roderick Chisholm leyó la primera ponencia, «Sucesos y proposiciones». Este artículo y el ensayo 9 se publicaron por primera vez en *Noûs* 4, 1970, de Wayne State University, Detroit, Michigan. Poco después, Chisholm me respondió en «States of Affairs Again», *Noûs* 5, 1971,

y yo le respondí nuevamente con el ensayo 10 en el mismo volumen de *Noûs*.

El ensayo 11, «Sucesos mentales», fue una de una serie de conferencias que impartieron varios filósofos en la Universidad de Massachusetts en 1968-1969. Estas conferencias se publicaron en Lawrence Foster y J.W. Swanson (comps.), *Experience and Theory*, The University of Massachusetts Press y Duckworth, 1970, y se reimprimen con autorización de los editores.

El ensayo 12, «La psicología como filosofía», se presentó en un simposio sobre la filosofía de la psicología en la Universidad de Kent en 1971. El artículo se publicó, con comentarios y respuestas, en S.C. Brown (comp.), *Philosophy of Psychology*, The Macmillan Press and Barnes, Noble, Inc., 1974, y se reimprime con autorización de Macmillan, Londres y Basingstoke, y de Barnes & Noble Books, Nueva York.

El ensayo 13, «La mente material», fue escrito en un café de Viena y presentado en la Sección 9, Metodología y Filosofía de las Ciencias Psicológicas, en el Cuarto Congreso Internacional de Lógica, Metodología y Filosofía de la Ciencia, en Bucarest, 1971. Las Actas del Congreso, compiladas por P. Suppes, L. Henkin, G.C. Moisil y A. Joja, se publicaron en la North-Holland Publishing Company en 1973, y mi artículo se reimprime con autorización de los editores.

El ensayo 14, «Hempel y la explicación de la acción», se leyó en un homenaje en honor del profesor Carl G. Hempel en la Universidad de Princeton en noviembre de 1975, y se publicó en *Erkenntnis* 10, 1976, pp. 239-253, *copyright* ©1976 por D. Reidel Publishing Company, Dordrecht, Holanda. Se reimprime con la autorización de D. Reidel Publishing Company.

«La teoría cognoscitiva del orgullo de Hume», el ensayo 15, se presentó en un simposio durante las reuniones de la Eastern Division de la American Philosophical Association en diciembre de 1976, y se publicó en el *Journal of Philosophy* 73, 1976. Se reimprime con autorización de los editores. Annette Baier y Keith Donnellan hicieron comentarios a mi ponencia en el simposio, y

Annette Baier amplió después sus comentarios y los convirtió en el artículo «Hume's Analysis of Pride», *Journal of Philosophy* 75, 1978.

Los ensayos 1, 3, 6, 7, 8, 11 y 12 se han publicado anteriormente en varios lugares e idiomas.

## INTRODUCCIÓN

Todos los ensayos de este libro se han publicado en otros lados, y cada uno ha sido diseñado de manera que pueda leerse con cierta autonomía. Sin embargo, aun cuando se han redactado a lo largo de una docena de años, existe una unidad en el tema y en la tesis general. El tema es el papel de los conceptos causales en la descripción y la explicación de la acción humana. La tesis es que la noción ordinaria de causa que se introduce en las explicaciones, tanto científicas como del sentido común, de asuntos no psicológicos, es esencial también para entender qué es actuar por una razón, tener una cierta intención al actuar, ser un agente, actuar en contra del propio mejor juicio o actuar libremente. La causa es el cemento del universo; el concepto de causa es lo que mantiene unida nuestra imagen del universo, una imagen que de otra manera se desintegraría en un dúptico de lo mental y lo físico.

Dentro de las tres grandes subdivisiones que he impuesto a los ensayos, el orden de publicación proporciona un esquema de organización bastante natural. Una cosa llevó a la otra, de las soluciones de un artículo surgieron los problemas del siguiente. Sin embargo, como será claro aun para el lector más benévolo, no todo resultó fácil; algunos problemas posteriores frecuentemente inspiraron revisiones drásticas de doctrinas anteriores. La unidad de la tesis general se presenta, en estas páginas, con una considerable inconsistencia diacrónica.

No se ha hecho ningún intento de ocultar las discrepancias entre propuestas anteriores y posteriores. Se han eliminado algunos errores accidentales y fealdades estilísticas y se han reducido las re-

dundancias. Hay todavía bastante redundancia, pero las propuestas en las que más insisto son generalmente aquellas que me causaron mayores problemas, así que espero que se encuentren enseñanzas o interés en aquello que pudiera parecer, y que probablemente pretendía ser, mera repetición. Otra razón por la que he dejado mis primeros pensamientos sin un cambio sustancial es que en el curso de los años han provocado comentarios y críticas, y sería de mala fe tratar de mover el blanco fuera de alcance cuando ya se ha disparado el tiro.

Éstas son algunas de las conexiones entre los ensayos:

*Ensayo 1.* «Acciones, razones y causas» fue una reacción en contra de la doctrina ampliamente aceptada de que la explicación de una acción intencional en términos de sus motivos o razones no podía relacionar a las razones y acciones como causa y efecto. Un argumento central era que las relaciones causales son esencialmente nomológicas y están basadas en la inducción, mientras que nuestro conocimiento de que un agente ha actuado por ciertas razones no depende usualmente de la inducción o del conocimiento de leyes serias. Este argumento había tenido una expresión influyente, aunque breve, en los *Libros azul y marrón* de Wittgenstein, que circularon ampliamente desde mediados de los años treinta (aunque se publicaron hasta 1958). En el ensayo 1 acepto la tesis de que las explicaciones teleológicas de la acción difieren de las explicaciones en las ciencias naturales en cuanto a que las primeras no incluyen esencialmente leyes, pero sostengo que ambos tipos de explicación pueden, y muchas veces deben, invocar conexiones causales.

*Ensayo 2.* La explicación de una acción intencional en términos de razones difiere de la explicación en las ciencias naturales en otro punto crucial: los contenidos proposicionales de las actitudes y de las creencias explicativas deben mantener cierta relación lógica con la descripción bajo la cual se explica la acción —una descripción que nos proporciona una intención con la que se realizó la acción. Pero ¿cuál es esta relación lógica? En el ensayo 2, «¿Cómo es posible la debilidad de la voluntad?», asumo que no es aceptable ninguna respuesta que frustre una explicación coherente del conflicto moral, de la debilidad de la voluntad y de otras formas de acción intencional, aunque sean irracionales. En este ensayo objeto la tesis, expuesta en el ensayo 1, de que las expresiones proposicionales de las razones



para una acción estén deductivamente relacionadas con la proposición que corresponde a la acción tal como la explican esas razones.

Las acciones intencionales que son contrarias al mejor juicio del actor son un desafío para las teorías causales de la acción. Porque, si las razones son causas, es natural suponer que las razones más fuertes son las causas más fuertes. Defiendo la tesis causal en el ensayo 2, argumentando que una razón que es la más fuerte causalmente no necesita ser la razón que el actor considera que proporciona los motivos más fuertes (mejores) para actuar.

*Ensayo 3.* Este ensayo, sobre la actuación [*agency*], se pregunta acerca de cuál es la relación entre un agente y un suceso que hace de ese suceso una acción. Se rechazan algunas propuestas y se plantean otras para reducir algunos casos a otros: de las acciones que no son intencionales se afirma que lo son bajo otras descripciones, y se argumenta que las acciones que aparentan incluir consecuencias de acciones son idénticas a la acción causante. Pero la demanda fundamental de análisis de la actuación (y por consiguiente, según el argumento, de la acción) queda sin respuesta.

*Ensayo 4.* Las teorías causales siempre han sido vulnerables a la crítica en el sentido de que no pueden ofrecer una explicación de la acción libre, o del ser libres para realizar un acto. En este ensayo trato de deshacer algunos de los argumentos que parecían mostrar que la libertad para actuar no puede ser un poder causal. Pero también critico algunos intentos de analizar ese poder y concluyo que, aun cuando la libertad para actuar es un poder causal, no puede analizarse ni definirse, por lo menos no sin apelar a la noción de intención. Debido a la íntima conexión que existe entre la libertad para actuar y la acción intencional, la conclusión de este ensayo contradice una nota optimista en el ensayo 1 que anticipaba, pero no ofrecía, una enunciación de condiciones suficientes para la acción intencional (libre).

*Ensayo 5.* Cuando escribí el ensayo 1 creía que, de los tres usos principales del concepto de intención que distingue Anscombe (actuar con una intención, actuar intencionalmente y tener la intención de actuar), el primero era el más básico. Actuar intencionalmente, argüí en el ensayo 1, sólo era actuar con alguna intención. Faltaba tener la intención, que yo pensaba que, de algún modo, podría entenderse fácilmente en términos de los otros. Me equivoqué. Cuando

pude por fin dedicarme a ello, me di cuenta de que era el más difícil de los tres; en contra de mi idea original, era, al parecer, la noción básica de la que dependen las otras; y lo que logré avanzar con respecto a esta noción socava en parte un tema importante del ensayo 1 —que «la intención con la que se hizo una acción» no se refiere a una entidad o estado de alguna clase.

*Ensayo 6.* «La forma lógica de las oraciones de acción» introduce un grupo de cinco artículos sobre la semántica de las oraciones que contienen verbos de acción, o de cambio en general, y de otras oraciones estrechamente relacionadas que contienen frases nominales que aparentemente se refieren a sucesos o acciones.

En el ensayo 1 me referí de un modo demasiado ligero a acciones bajo una descripción, o a dos descripciones de la misma acción; pero me preocupaba el hecho de que la mayoría de las oraciones que conciernen a acciones no contenían una descripción u otro recurso que sirviera para referirse a un suceso o acción. Algunos filósofos han considerado esta incursión en asuntos de forma lógica como ajena a la teoría de la acción. Esta actitud es, sin embargo, riesgosa, pues es claro, creo yo, que muchas teorías de la acción que en un principio parecen atractivas revelan confusiones subyacentes cuando se las somete a un escrutinio formal cuidadoso. Cuando escribí el ensayo 1, no se me había ocurrido todavía que no debemos considerar que una oración como «Eva comió la manzana» contiene una referencia singular a un suceso; su forma lógica es diferente a la de «La ingestión de la manzana por Eva ocurrió», aunque la segunda implica lógicamente la primera.

He anexado al ensayo 6 algunas réplicas a varias sugerencias y críticas que son pertinentes a la tesis que ahí se sostiene.

El *ensayo 7* aplica a la causalidad la lección del ensayo 6. El ensayo 1 dependía en gran medida del hecho de que los sucesos pueden describirse en formas lógicamente independientes, de modo tal que las razones y las acciones, descritas tal como deben serlo para resaltar la intención con la que se realizó la acción, no pueden describirse como haría falta para que apuntaran a leyes causales serias. No obstante, las razones y las acciones pueden ser causa y efecto. Sin embargo, en el ensayo 1 no tenía todavía una idea clara de lo que quería decir con la descripción de un suceso y, por

consiguiente, tampoco tenía idea de cuál pudiese ser la forma lógica de las oraciones causales singulares.

*Ensayo 8.* La doctrina del ensayo 6 postulaba que ciertas descripciones de acciones o sucesos, si es que refieren a algo, han de referir a la misma cosa. En ese caso, la mayoría de los modificadores adverbiales no pueden cambiar la referencia de la expresión que modifican, y (como es de esperar —aunque algunas teorías actuales lo objeten) el sustituir términos singulares correferenciales en la descripción de una acción o suceso no puede cambiar (excepto en ciertos contextos) la acción o suceso al que se refieren.

Pero estos casos no tocan otros problemas más interesantes con respecto a la individuación de los sucesos, aquellos en los que la lógica no basta para decidir el asunto. El ensayo 8 se ocupa en gran medida de otros criterios de individuación. En particular, introdujo por primera vez en letra impresa (hasta donde yo lo sé) un enigma que, desde entonces, ha sido muy discutido: si *A* mata a *B* envenenándolo, parece difícil negar que tiene que haber un asesinato que es idéntico a un envenenamiento. Pero un asesinato implica una muerte, y la muerte puede ocurrir mucho tiempo después del acto que la causó. ¿Cómo puede entonces el envenenamiento ser idéntico al asesinato? El ensayo 8 propuso una respuesta que ha sido ampliamente criticada y defendida. La respuesta todavía me parece mejor que las alternativas radicales, aunque ahora pienso que se podría mejorar en ciertos aspectos.

Los *ensayos 9 y 10* defienden y elaboran la ontología de sucesos que se propone en el ensayo 6. Ambos ensayos fueron escritos en respuesta a artículos de Roderick Chisholm.

Los siguientes cuatro artículos exploran el papel de las leyes en la explicación de las acciones y de otros fenómenos psicológicos.

El *ensayo 11* argumenta que, aun cuando los fenómenos psicológicos y físicos se conectan causalmente y esto implica lógicamente que existen leyes estrictas que gobiernan a los sucesos que las ejemplifican, no hay, sin embargo, leyes estrictas que cubran a los sucesos o estados descritos en términos psicológicos. Al explicar cómo puede ocurrir esto surge una versión de la teoría de la identidad que yo llamo *monismo anómalo*.

En el *ensayo 12* hay un mayor desarrollo de los argumentos en contra de la posibilidad de leyes psicofísicas estrictas y se pone énfasis

sis en la importancia fundamental que tiene un concepto normativo de racionalidad en la atribución de creencias, intenciones, deseos y otras actitudes de esta clase.

El *ensayo 13* plantea la cuestión acerca de cuánto podemos esperar aprender sobre la psicología del pensamiento y de la acción de los avances en la neuropsicología, y llega a conclusiones que están íntimamente relacionadas con las de los ensayos 11 y 12.

El *ensayo 14*, al igual que el ensayo 8, lo escribí para celebrar las virtudes, tanto intelectuales como de otro tipo, de Carl Hempel. En este ensayo, «Hempel y la explicación de la acción», se lo reconoce como un antiguo partidario de la teoría causal de la acción, pero se disiente respecto al papel que le atribuye a las leyes empíricas en la explicación de acciones.

El *ensayo 15*, sobre la teoría del orgullo de Hume, defiende un análisis humeano de ciertas pasiones y emociones y, en particular, defiende la tesis de Hume de que la creencia de que uno tiene una casa bella (por ejemplo) es una condición causal de estar orgulloso de que uno tiene una casa bella. Al igual que en el ensayo 1 y en los subsiguientes, alego que podría ser una verdad necesaria el que un suceso o estado, descrito de cierta manera, tenga cierta causa. El que esto sea consistente o no con el análisis que Hume hace de la causalidad es otra cuestión y, en mi opinión, es una cuestión abierta. En cualquier caso, si es correcta la tesis que defiendo aquí acerca de la naturaleza del orgullo «proposicional» y de otras emociones similares, esto muestra cómo el orgullo, al igual que las acciones para las cuales podría ofrecer una razón, forma parte de las cadenas causales que ayudan a explicar, y en cierta medida a justificar, patrones humanos de sentimiento, pensamiento y conducta.

Debo dar gracias a muchas personas, entre ellas y en particular a Max Black, Michael Bratman, Joel Feinberg, Paul Grice, Stuart Hampshire, Gilbert Harman, Carl Hempel, Merrill Hintikka, Georg Kreisel, Sue Larson, David Lewis, Harry Lewis, Mary Mothersill, David Nivison, David Pears, Richard Reiss, Richard Rorty, Allison Ryan, David Sachs, J.J. Smart, P.F. Strawson, Patrick Suppes, Irving Thalberg, David Wiggins y Kathleen Wilkes. Akeel Bilgrami preparó la bibliografía del libro e hizo mucho para mejorar su estilo y contenido. Con Sue Larson colaboró en la elaboración del índice analítico. Sin su ayuda y estímulo no habría podido hacer este libro.

La National Science Foundation y la Guggenheim Foundation apoyaron la elaboración de algunos de los ensayos. Dos de ellos fueron escritos durante el año que pasé en el Center for Advanced Study in the Behavioral Sciences de Stanford, y otro durante mi visita al All Souls College en Oxford.

Las Conferencias Gavin David Young, que impartí en la Universidad de Adelaida en 1968, se basaron en gran medida en borradores de ensayos que se publican aquí. Mi agradecimiento a las personas que hicieron comentarios a esas conferencias y, en especial, a Jack Smart, cuyas tesis acerca de la relación entre lo mental y lo físico, aunque en ciertos aspectos diametralmente opuestas a las mías, constituyeron un gran estímulo para mi propio pensamiento. También le agradezco que me señalara un error astronómico (no de tamaño, sino de tema).

La mayoría de las ideas en este libro las elaboré y las probé en discusiones con aquellos a los que yo, en algún momento, estaba supuestamente enseñando. Daniel Bennett me animó mucho a empezar con este tema. Su tesis, en teoría escrita bajo mi dirección, me abrió los ojos hacia lo que estaba sucediendo a mediados de los cincuenta en Oxford, y me influyó nuevamente cuando, siendo ambos colegas en la Universidad de Stanford, dimos al alimón un seminario sobre la teoría de la acción.

John Wallace y yo hablamos interminablemente acerca de problemas de filosofía del lenguaje, primero cuando él era estudiante en Stanford, luego durante un año en Atenas, Corfú y Gründelwald y, más tarde, cuando fuimos colegas en las universidades de Princeton y Rockefeller. Las ideas clave de los ensayos 6 y 7 surgieron de estas conversaciones.

En 1968 Nancy Hirschberg me invitó a dar una charla en el Departamento de Psicología de la Universidad de Illinois en Champaign; el ensayo 11 fue uno de los resultados. En la composición de los últimos cinco ensayos me guiaron su conocimiento de la psicología y sus consejos justos. Su entusiasmo, alegría y cariño hicieron divertido el trabajo y maravillosa la diversión.

# **INTENCIÓN Y ACCIÓN**

## 1. ACCIONES, RAZONES Y CAUSAS

¿Qué relación hay entre razón y acción cuando una razón explica una acción ofreciendo la razón del agente para hacer lo que hizo? Podemos llamar *racionalizaciones* a tales explicaciones y decir que la razón *racionaliza* la acción.

En este ensayo quiero defender la posición antigua —y de sentido común— según la cual la racionalización es una especie de explicación causal. Esta defensa requiere, sin duda, de alguna reelaboración, pero no parece necesario abandonar totalmente dicha posición, tal como lo pretenden muchos autores recientes.<sup>1</sup>

### I

Una razón racionaliza una acción sólo si nos lleva a ver algo que el agente vio, o pensó ver, en su acción; algún rasgo, consecuencia o aspecto de la acción que el agente quiso, deseó, apreció; que le pareció atractivo, benéfico, obligatorio, agradable, o que consideró como su deber. No podemos explicar por qué alguien hizo lo que hizo diciendo simplemente que esa acción particular le pareció atractiva: tenemos que señalar qué fue lo que le pareció atractivo de la acción. Por lo tanto, siempre que alguien hace algo por una razón, puede

1. Algunos ejemplos: Gilbert Ryle, *The Concept of Mind*; G.E.M. Anscombe, *Intention*; Stuart Hampshire, *Thought and Action*; H.L.A. Hart y A.M. Honoré, *Causation in the Law*; William Dray, *Laws and Explanation in History* y casi todos los libros de la serie editada por R.F. Holland, *Studies in Philosophical Psychology*, comprendida la obra de Anthony Kenny, *Action, Emotion and Will* y la de A.I. Melden, *Free Action*. Los números entre paréntesis se refieren a páginas de estas obras.

caracterizarse: (a) como si tuviera algún tipo de actitud favorable hacia acciones de una clase determinada, y (b) como si creyera (o supiera, percibiera, notara, recordara) que su acción es de esa clase. Deben incluirse en (a) actitudes tales como deseos, impulsos, instintos y una gran variedad de convicciones morales, principios estéticos, prejuicios económicos, convencionalismos sociales, metas y valores públicos y privados, en la medida en que éstos puedan interpretarse como actitudes del agente dirigidas a cierta clase de acciones.

La palabra «actitud» vale aquí para todo, porque debe abarcar no sólo los rasgos permanentes de carácter que se muestran en la conducta de alguien a lo largo de toda su vida, como el amor a los niños o la afición a compañeros escandalosos, sino también el capricho más efímero que impulsa a una acción única, como el deseo súbito de tocar el hombro de una mujer. En general, las actitudes favorables no deben confundirse con las convicciones, por pasajeras que sean, de que toda acción de cierta clase debería realizarse, de que vale la pena realizarla o de que, hechas todas las consideraciones, es deseable. Por el contrario, un hombre puede tener ganas durante toda su vida de beberse un bote de pintura, sin que nunca, ni siquiera en el momento de ceder, piense que valdría la pena hacerlo.

A menudo, dar la razón por la cual un agente hizo algo consiste en nombrar la actitud favorable (a) o la creencia relacionada (b), o ambas; permítaseme llamar a este par la *razón primaria* por la que el agente realizó la acción. Ahora es posible reformular la afirmación de que las racionalizaciones son explicaciones causales y también dotar de una estructura al argumento enunciando dos tesis sobre las razones primarias:

1. Para entender cómo una razón de cualquier tipo racionaliza una acción, es necesario y suficiente que veamos, por lo menos en sus rasgos esenciales, cómo construir una razón primaria.
2. La razón primaria de una acción es su causa.

A continuación ofreceré consecutivamente argumentos a favor de estas dos tesis.



## II

Presiono el interruptor, enciendo la luz e ilumino el cuarto. Sin saberlo, también alerto a un merodeador de mi presencia en la casa. En este caso no tuve que hacer cuatro cosas sino una sola, de la cual se han dado cuatro descripciones.<sup>2</sup> Presioné el interruptor porque quería encender la luz, y al decir que quería encender la luz explico (doy mi razón, racionalizo) el acto de presionar el interruptor. Pero al dar esta razón no racionalizo que haya alertado al merodeador ni iluminado el cuarto. Dado que las razones pueden racionalizar lo que alguien hace cuando se lo describe de cierta manera y no cuando se lo describe de otra, no podemos considerar lo que se hizo simplemente como un término en oraciones como «la razón para presionar el interruptor fue que quería encender la luz»; de otro modo, nos veríamos obligados a concluir, a partir del hecho de que presionar el interruptor es idéntico a alertar al merodeador, que mi razón para alertar al merodeador era que quería encender la luz. Para señalar este carácter cuasi-intensional<sup>3</sup> de las descripciones de

2. Podríamos no llamar una acción al hecho de que yo, no intencionalmente, advirtiera de mi presencia al merodeador, pero de esto no debería inferirse que advertir al merodeador fuera, por lo tanto, algo distinto de presionar el interruptor, digamos, sólo su consecuencia. Las acciones, las realizaciones y los sucesos que no suponen una intención se parecen entre sí en tanto que frecuentemente nos referimos a ellos o se los define en términos de alguna etapa terminal, de un resultado o consecuencia. La palabra «acción» no figura con mucha frecuencia en el lenguaje ordinario, y cuando lo hace se la reserva usualmente para ocasiones bastante portentosas. Al llamar acción a todo lo que un agente hace intencionalmente incluyendo las omisiones intencionales, sigo una útil costumbre filosófica. Lo que se necesita en realidad es algún término adecuadamente genérico para llenar esta laguna: supongamos que «A» es una descripción de una acción, «B» es una descripción de algo hecho voluntaria pero no intencionalmente, y «C» es una descripción de algo hecho involuntariamente y de manera no intencional. Finalmente, supongamos que  $A = B = C$ . Entonces A, B y C son una misma —¿qué? «Acción», «suceso», «cosa hecha»; cada una de estas expresiones tiene un halo extraño, por lo menos en algunos contextos, cuando se une a la descripción equivocada. Sólo la pregunta: «¿Por qué hiciste (hizo) A?» tiene la generalidad que realmente se requiere. Obviamente, el problema se agrava seriamente si suponemos, como lo hace Melden, que una acción («levantar uno el brazo») puede ser idéntica a un movimiento corporal («que el brazo de uno se levante»).

3. «Cuasi-intensional» porque, además de tener un aspecto intensional, la descripción de la acción también debe tener una referencia en las racionalizaciones;

acciones en las racionalizaciones, hay que enunciar, con un poco más de precisión, una condición necesaria de las razones primarias:

C 1. *R* es una razón primaria por la que un agente realizó la acción *A* en la descripción *d*, sólo si *R* consiste en una actitud favorable del agente hacia las acciones que poseen cierta propiedad y en una creencia suya de que *A* en la descripción *d* tiene esa propiedad.

¿Cómo puede ser mi querer encender la luz (parte integrante de) una razón primaria si aparentemente carece del elemento de generalidad que se requiere? Nos puede engañar el paralelismo verbal entre «encendí la luz» y «quise encender la luz». Lo primero se refiere claramente a un suceso particular, y por eso concluimos que lo segundo tiene por objeto ese mismo suceso. Es obvio, por supuesto, que las dos oraciones no pueden referirse de la misma manera al suceso que consiste en mi encender la luz, ya que la verdad de «encendí la luz» requiere de la existencia del suceso, y no así la verdad de «quise encender la luz». Si la referencia fuera la misma en ambos casos, la segunda oración implicaría lógicamente la primera: pero de hecho las dos oraciones son lógicamente independientes. Lo que es menos obvio, por lo menos mientras no le prestemos atención, es que el suceso que al ocurrir hace verdadera «encendí la luz» no puede ser llamado objeto, por intencional que sea, de «quise encender la luz». Si encendí la luz, tuve que hacerlo en un momento preciso y de una manera determinada; todos los detalles están fijados. Pero no tiene sentido exigir que mi deseo se dirija a una acción realizada en cierto momento y hecha de alguna manera única. Cualquier acción de un conjunto indefinidamente grande satisfaría el deseo y podría considerarse igualmente elegible como su objeto. A menudo el querer y el desear se dirigen a objetos físicos. Sin embargo, «quiero ese reloj de oro que está en la vitrina» no es una razón primaria y sólo porque sugiere una razón primaria explica por qué entré en la tienda: por ejemplo, sugiere que quería comprar el reloj.

de otro modo podría ser verdadero que se hubiera hecho una acción por cierta razón y que, sin embargo, no se hubiera realizado la acción. Compárese «el autor de *Waverley*» en «Jorge IV sabía que el autor de *Waverley* escribió *Waverley*». Este rasgo semántico de las descripciones de acciones se discute más ampliamente en los ensayos 3 y 6.

En vista de que «quise encender la luz» y «encendí la luz» son lógicamente independientes, la primera puede usarse para dar una razón de la verdad de la segunda. Esta razón ofrece una información mínima: implica que la acción fue intencional —y querer tiende a excluir algunas de las otras actitudes favorables, tales como el sentido del deber o de la obligación. Pero la exclusión depende mucho de la acción y del contexto de la explicación. Querer nos parece un término pálido comparado con ansiar; sin embargo, sonaría extraño negar que un hombre que ansiara una mujer o una taza de café no las quisiera. En realidad, no deja de ser natural tratar el querer como un género que incluye todas las actitudes favorables como especies suyas. Cuando se considera así y cuando sabemos que alguna acción es intencional, es fácil responder a la pregunta «¿por qué lo hiciste?» contestando «pues por ninguna razón», queriendo decir no que no haya ninguna razón, sino que no hay ninguna razón *ulterior*, ninguna razón que no pueda inferirse del hecho de que la acción se hizo intencionalmente; en otras palabras, ninguna razón aparte del querer hacerla. Este último punto no es esencial a la presente discusión, pero no carece de interés pues defiende la posibilidad de definir una acción intencional como una acción que se hace por una razón.

Una razón primaria consiste en una creencia y en una actitud, pero generalmente es ocioso mencionar las dos. Si me dices que estás soltando el foque porque piensas que eso evitará que se vaya hacia atrás la vela mayor, no necesito que me digas además que quieres detener el desplazamiento hacia atrás de la vela mayor, y si dices que me estás haciendo una seña con los dedos porque quieres insultarme, no tiene objeto añadir que piensas que al hacerme esa seña con los dedos vas a insultarme. De manera semejante, muchas explicaciones de las acciones en términos de razones que no son primarias no requieren que se mencione la razón primaria para completarlas. Si digo que arranco las malas yerbas porque quiero que el césped esté bonito, sería fatuo añadir lo siguiente: «Y, por tanto, considero que hay algo deseable en cualquier acción que embellezca el césped o tenga buenas probabilidades de embellecerlo». ¿Por qué insistir en que hay algún *paso*, lógico o psicológico, cuando se transfiere el deseo de un fin que no es una acción a las acciones consideradas como medios? Apoya igualmente el razonamiento que el fin deseado explica la acción sólo si el agente desea lo que concibe como medio.

Afortunadamente no es necesario clasificar y analizar la gran variedad de emociones, sentimientos, estados de ánimo, motivos, pasiones y apetitos cuya mención puede responder a la pregunta «¿por qué lo hiciste?» para ver cómo se incluye una razón primaria cuando las expresiones mencionadas racionalizan la acción. La claustrofobia ofrece una razón por la que un hombre abandona una fiesta porque sabemos que la gente quiere evitar lo que teme, evadirlo, ponerse a salvo y alejarse de ello. Los celos son el motivo de un envenenamiento, entre otras cosas porque el envenenador cree que su acción dañará a su rival, suprimirá la causa de su agonía o reparará una injusticia, y éstas son las cosas que quiere hacer un hombre celoso. Cuando nos enteramos de que un hombre timó a su hijo por codicia, no necesariamente sabemos cuál fue la razón primaria, pero sí sabemos que la hubo y cuál era su naturaleza general. Ryle analiza la frase «se jactó por vanidad» como «se jactó al toparse con el desconocido y el actuar así satisface la proposición legaliforme de que siempre que vea una oportunidad de granjearse la admiración y la envidia de otros, hace cualquier cosa que considere capaz de producir esta admiración y esta envidia» (p. 89). Este análisis se critica con frecuencia, y quizá no injustamente, por la razón de que un hombre puede jactarse por vanidad una sola vez. Pero si el jactancioso de Ryle hizo lo que hizo por vanidad, entonces es verdad algo que implica lógicamente el análisis de Ryle: el jactancioso quería granjearse la admiración y la envidia de los demás y creía que su acción produciría esta admiración y esta envidia; verdadero o falso, el análisis de Ryle no prescinde de las razones primarias sino que depende de ellas.

Conocer una razón primaria por la cual alguien actuó de cierta manera es conocer una intención con la que se hizo la acción. Si doy vuelta a la izquierda en la bifurcación porque quiero ir a Katmandú, mi intención al dar vuelta a la izquierda es ir a Katmandú. Pero conocer la intención no necesariamente es conocer la razón primaria en todo su detalle. Si Jaime va a la iglesia con la intención de complacer a su madre, debe de tener alguna actitud favorable hacia el agrado que produce a su madre, pero se necesita más información para decir si su razón consiste en que goza al agradar a su madre o en que lo considera correcto o lo concibe como su deber o como una obligación. La expresión «la intención con la que Jaime fue a la

iglesia» tiene la forma aparente de una descripción, pero de hecho es sincategoremática y no se la puede tomar como si se refiriera a una entidad, estado, disposición o suceso. Su función en un contexto es la de generar nuevas descripciones de acciones en términos de sus razones: así, «Jaime fue a la iglesia con la intención de agrandar a su madre» da lugar a una descripción nueva y más completa de la acción descrita como «Jaime fue a la iglesia». Se da en esencia el mismo proceso cuando respondo a la pregunta «¿por qué te mueves de esa manera?», con «estoy hilando, tejiendo, bordando, haciendo ejercicio, remando, acariciando, adiestrando pulgas».

A menudo, describir directamente el resultado pretendido explica mejor una acción que afirmar que el resultado era objeto de una intención o de un deseo. Decir «te calmará los nervios» explica por qué te sirvo un trago tan eficazmente como «quiero hacer algo para calmarte los nervios», porque lo primero en el contexto de la explicación implica lo segundo; pero lo primero lo hace mejor porque, de ser verdadero, los hechos justificarán la acción que yo elija. En vista de que justificar y explicar una acción van unidas de manera tan frecuente, a menudo indicamos la razón primaria de una acción aseverando algo que, de ser verdadero, también verificaría, reivindicaría o apoyaría la creencia o actitud pertinentes del agente. «Sabía que debía devolverlo», «el periódico decía que nevaría», «pisaste *mis* dedos», son todas expresiones tales que, en los contextos apropiados de ofrecer razones, realizan esa doble función que nos es familiar.

Según esta interpretación, el papel justificativo de una razón depende del papel explicativo, pero no viceversa. El que hayas pisado mis dedos ni explica ni justifica que yo pise los tuyos, a menos que yo crea que pisaste mis dedos, y la creencia por sí sola, sea verdadera o falsa, explica mi acción.

### III

A la luz de una razón primaria una acción se revela como algo coherente con ciertos rasgos del agente, sean estos efímeros o duraderos, característicos o no; y se muestra al agente en su papel de Animal Racional. Siempre es posible construir (con un poco de ingenio) un silogismo cuyas premisas correspondan a la creencia y

la actitud de la razón primaria de una acción, y de las cuales se sigue que la acción tiene alguna «característica de deseabilidad» (como la llama Anscombe).<sup>4</sup> Así, hay cierto sentido irreductible —aunque un poco anémico— según el cual toda racionalización justifica: desde el punto de vista del agente en el momento en que actuó, había algo que decir a favor de la acción.

Al advertir que las explicaciones causales no teleológicas no revelan el elemento de justificación que aportan las razones, algunos filósofos han concluido que el concepto de causa que se aplica en otras partes no puede aplicarse a la relación entre razones y acciones y que el patrón de justificación proporciona, en el caso de las razones, la explicación requerida. Pero supongamos que se concede que las razones solas justifican las acciones al explicarlas: de ello no se sigue que la explicación no sea también —y necesariamente— causal. De hecho la primera condición de las razones primarias C 1 tiene como propósito ayudar a separar las racionalizaciones de otras clases de explicación. Si, como quiero argumentar, la racionalización es una especie de explicación causal, entonces la justificación, en el sentido dado en C 1, es por lo menos una propiedad diferenciadora. ¿Qué debe decirse de la otra tesis según la cual la justificación es un tipo de explicación tal que no necesita recurrir a la noción ordinaria de causa? Aquí es necesario decidir qué se está incluyendo bajo el término justificación. Podría interpretarse como si abarcara sólo lo que está dado en C 1, a saber, que el agente tiene ciertas creencias y actitudes a la luz de las cuales la acción es razonable. Pero entonces ciertamente se ha dejado fuera algo esencial, porque alguien puede tener una razón para una acción y, sin embargo, realizar la acción sin que esta razón sea la razón por la que la hizo. La idea de que el agente realizó una acción *porque* tenía una razón es fundamental para la relación entre una razón y la acción que explica. Por supuesto,

4. Anscombe niega que el silogismo práctico sea deductivo. Lo hace en parte porque considera, como Aristóteles, que el silogismo corresponde a un razonamiento práctico (mientras que para mí es sólo una parte del análisis del concepto de una razón por la que alguien actuó) y, por lo tanto, se ve obligada, siguiendo también en esto a Aristóteles, a pensar que la conclusión del silogismo práctico corresponde a un juicio, no sólo que la acción tiene una característica deseable, sino que ella misma es deseable (es razonable, digna de hacerse, etc.). El razonamiento práctico se discute ampliamente en el ensayo 2.

también podemos incluir esta idea en la justificación; pero entonces, hasta no poder dar cuenta de la fuerza de ese «porque», la noción de justificación será tan oscura como la noción de razón.

Cuando preguntamos por qué alguien actuó como lo hizo veremos que se nos dé una interpretación. Su conducta nos parece extraña, rara, extravagante, carente de sentido, inapropiada, incoherente; o quizá ni siquiera reconozcamos en ella una acción. Cuando nos enteramos de su razón tenemos una interpretación, una descripción nueva de lo que hizo, que lo hace encajar dentro de un modo familiar de ver las cosas. Este modo de ver las cosas comprende algunas de las creencias y actitudes del agente; quizá también algunas de sus metas, fines, principios, rasgos generales de carácter, virtudes o vicios. Además de esto, la redesccripción de una acción que ofrece una razón puede colocar a la acción en un contexto social, económico, lingüístico o valorativo más amplio. Enterarse mediante el conocimiento de la razón de que el agente concebía su acción como una mentira, como el pago de una deuda, como un insulto, como el cumplimiento de una obligación familiar o como un gambito de caballo, es captar el propósito de la acción en su contexto de reglas, costumbres, convenciones y expectativas.

Observaciones como éstas, inspiradas en el último Wittgenstein, han sido elaboradas con sutileza y penetración por cierto número de filósofos. Y no podemos negar que es verdad lo siguiente: cuando explicamos una acción dando una razón, de hecho redescibimos la acción; redescibir la acción le asigna a ésta un lugar en un patrón y, de esta manera, se explica la acción. Surge aquí la tentación de derivar dos conclusiones que no se siguen. En primer lugar, a partir del hecho de que al dar razones meramente redescibimos la acción y de que las causas son distintas de los efectos, no podemos inferir que, por lo tanto, las razones no son causas. Las razones, al ser creencias y actitudes, ciertamente no son idénticas a las acciones; pero, lo que es más importante, a los sucesos se los redescibe, con frecuencia, en términos de sus causas. (Supóngase que alguien fue lastimado. Podríamos redescibir este suceso «en términos de una causa» diciendo que fue quemado.) En segundo lugar, es un error pensar que sólo porque situar la acción en un patrón más amplio es explicarla, entendemos ya de qué tipo de explicación se trata. Hablar de patrones y contextos no responde la pregunta acerca de cómo las

razones explican las acciones, ya que el patrón o contexto pertinente contiene a ambas, a la razón y a la acción. Una manera de explicar un suceso consiste en ubicarlo en el contexto de su causa; la causa y el efecto constituyen el tipo de patrón que explica el efecto, en un sentido de «explicar» que entendemos tan bien como cualquier otro. Si la razón y la acción ilustran un patrón distinto de explicación, tiene que identificarse ese patrón.

Permítaseme defender esta cuestión a propósito de un ejemplo de Melden. El conductor de un automóvil levanta el brazo para hacer una señal. Su intención, esto es, señalar, explica su acción, levantar el brazo, al describirla como hacer una señal. ¿Cuál es la pauta que explica la acción? ¿Se trata de la pauta familiar de una acción hecha por una razón? Entonces la pauta efectivamente explica la acción, pero sólo porque se supone la relación entre razón y acción que queremos analizar. ¿O se trata más bien de esta otra pauta?: el hombre conduce un automóvil, se está acercando a una esquina, sabe que debe hacer una señal, sabe cómo hacerlo, levantando el brazo. Y ahora, en este contexto, levanta el brazo. Es posible que, tal como lo sugiere Melden, si todo esto sucede, el conductor sí haga una señal. Y entonces la explicación sería la siguiente: si en estas condiciones un hombre levanta el brazo, entonces hace una señal. La dificultad consiste obviamente en que esta explicación no toca la cuestión de por qué levantó el brazo. Ese hombre tenía una razón para levantar el brazo, pero no se ha mostrado que ésta haya sido la razón por la que lo hizo. Si la descripción «hacer una señal» explica su acción ofreciendo su razón, entonces hacer la señal debe ser intencional; pero en la versión que acabamos de dar, puede no serlo.

Si, tal como pretende Melden, las explicaciones causales «no tienen absolutamente nada que ver con la comprensión que estamos buscando» de la acción humana (p. 184), entonces carecemos de un análisis del término «porque» en locuciones tales como «lo hizo porque. . . » en las que terminamos la frase nombrando una razón. Hampshire observa a propósito de la relación entre las razones y la acción: «En filosofía, esta conexión [ . . . ] sin duda debe parecerle a uno enteramente misteriosa» (p. 166). Hampshire rechaza el intento de Aristóteles de resolver el misterio introduciendo el concepto de deseo como un factor causal, sobre la base de que la teoría resultante es demasiado clara y definida para adecuarse a todos los casos,



y de que «sigue sin haber ninguna razón de peso para insistir en que la palabra “desear” *tenga* que formar parte de toda formulación completa de las razones para actuar» (p. 168). Convengo en que el concepto de deseo es demasiado estrecho, pero he sostenido que, por lo menos en un gran número de casos típicos, tiene que suponerse la presencia de alguna actitud favorable si la formulación de las razones que tuvo un agente para actuar ha de ser inteligible. Hampshire no ve cómo pueda considerarse como verdadero o falso el esquema de Aristóteles, «pues no queda claro cuál podría ser la base de la evaluación o qué tipo de prueba podría ser decisiva» (p. 167). Sin embargo, yo insistiría en que, a falta de una alternativa satisfactoria, el mejor argumento en favor de un esquema como el de Aristóteles es que sólo éste promete dar cuenta de la «conexión misteriosa» entre las razones y las acciones.

#### IV

Para convertir el primer «y» en un «porque» en la oración: «hacía ejercicio y quería adelgazar y pensaba que con el ejercicio lo lograría», como el paso fundamental<sup>5</sup> tenemos que complementar la condición C 1 de la siguiente manera:

**C 2 Una razón primaria de una acción es su causa.**

Ahora las consideraciones en favor de C 2 son obvias, según espero: en lo que resta de este ensayo deseo defender C 2 de varios ataques de diverso tipo y aclarar a lo largo de este proceso la noción de explicación causal que está en juego.

A. He aquí el primer tipo de ataque: las razones primarias consisten en actitudes y creencias que son estados o disposiciones, no sucesos; por tanto, no pueden ser causas.

Es fácil replicar que a los estados, a las disposiciones y a las condiciones se los denomina frecuentemente causas de los sucesos. Así, se dice que el puente se desplomó por un defecto estructural;

5. Digo «como el paso fundamental» para descartar cualquier sugerencia de que C 1 y C 2 son conjuntamente *suficientes* para definir la relación entre las razones y las acciones que explican. Para una discusión de este asunto véanse la Introducción y el ensayo 4.

que el avión se estrelló al despegar porque la temperatura del aire era anormalmente alta; que el plato se rompió porque estaba resquebrajado. Sin embargo, esta réplica no refuta una cuestión estrechamente relacionada. El mencionar una condición causal para un suceso nos proporciona una causa sólo si se supone que hubo también un suceso precedente. Pero ¿cuál es el suceso precedente que causa una acción?

En muchos casos no es difícil en absoluto encontrar sucesos estrechamente asociados a la razón primaria. Los estados y las disposiciones no son sucesos, pero sí lo es el arranque de un estado o disposición. Un deseo de lastimar tus sentimientos puede surgir en el momento en que me irritas; puedo empezar a querer comerme un melón justo en el momento de ver uno y las creencias pueden nacer en el momento en que notamos, percibimos, aprendemos o recordamos algo. Quienes han sostenido que no hay sucesos mentales adecuados para calificar como causas de las acciones frecuentemente han pasado por alto lo que es obvio, porque han insistido en que un suceso mental debe ser observado o notado (más bien que un observar o notar), o que tiene que ser como una estocada, un ataque de náusea, una punzada, una palpitación, un misterioso aguijón de la conciencia o un acto de la voluntad. Al discutir Melden acerca del conductor que alza el brazo para señalar que girará, reta a los que quieren explicar causalmente las acciones a que identifiquen «un suceso que sea peculiar y común a todos esos casos» (p. 87), quizá un motivo o una intención; en todo caso «algún sentimiento o experiencia particular» (p. 95). Pero evidentemente hay un suceso mental: en algún momento el conductor notó (o pensó notar) que había llegado a la esquina donde debía girar, y ése es el momento en que hizo la señal. Durante toda actividad duradera, como conducir un automóvil, o durante toda realización compleja, como cruzar a nado el Helesponto, hay propósitos, normas, deseos y hábitos más o menos fijos que dan una dirección y una forma a toda la empresa, y hay un suministro continuo de información sobre lo que se está haciendo y sobre los cambios del medio, de acuerdo con el cual regulamos y ajustamos nuestras acciones. Otorgarle importancia a la conciencia que tiene un conductor de que ha llegado el momento de girar llamándola una experiencia, o incluso un sentimiento, es sin duda exagerado; pero independientemente de que merezca un nom-

bre o no, más vale que sea la razón por la que el conductor levanta su brazo. En este caso, y por lo general, puede no haber nada a lo que podamos llamar un motivo, pero si mencionamos un propósito tan general como el de querer llegar a salvo a su destino, es claro que el motivo no es un suceso. La intención con la que el conductor alza el brazo tampoco es un suceso, porque no es en absoluto cosa alguna, ni un suceso, ni una actitud, ni una disposición, ni un objeto. Finalmente, Melden le pide al teórico causal que encuentre un suceso que sea peculiar y común a todos los casos en los que un hombre alza intencionalmente el brazo, y esto —tenemos que admitirlo— no puede hacerse. Pero entonces tampoco puede presentarse una causa única y común a los derrumbamientos de puentes, los accidentes de aviones y las rupturas de platos.

El conductor que hace la señal sí puede contestar la pregunta: «¿por qué levantaste el brazo cuando lo hiciste?», y por su respuesta nos enteramos del suceso que causó la acción. Pero ¿puede un agente responder siempre esa pregunta? A veces la respuesta aludirá a un suceso mental que no proporciona una razón, como el decir «finalmente me decidí». Sin embargo, también parece haber casos de acciones intencionales en los que no podemos explicar en absoluto por qué actuamos como lo hicimos. En tales casos, la explicación en términos de razones primarias es paralela a la explicación del desplome del puente por un defecto estructural: ignoramos el suceso o la secuencia de sucesos que condujeron al (causaron el) desplome, pero estamos seguros de que hubo tal suceso o secuencia de sucesos.

*B.* Según Melden, una causa tiene que ser «lógicamente distinta de su supuesto efecto» (p. 52); pero una razón de una acción no es lógicamente distinta de la acción; por lo tanto, las razones no son causas de las acciones.<sup>6</sup>

Ya se ha sugerido antes una de las formas que podría tener este argumento. Puesto que una razón hace inteligible una acción al describirla, no tenemos aquí dos sucesos, sino uno en diferentes

6. Puede encontrarse el argumento en una u otra versión en Kenny, Hampshire y Melden, así como en Peter Winch, *The Idea of a Social Science* y en R.S. Peters, *The Concept of Motivation*. Por supuesto, el tratamiento de los motivos que hace Ryle en *The Concept of Mind* inspiró el argumento en una de sus formas.

descripciones. Sin embargo, las relaciones causales exigen sucesos distintos.

Se puede caer en el error de pensar que mi presionar el interruptor fue la causa de que yo encendiera la luz (de hecho, fue la causa de que la luz se encendiera). Pero no se sigue de ello que sea un error considerar que «mi razón para presionar el interruptor fue que quería encender la luz» implique lógicamente, en parte, «presioné el interruptor y esta acción puede describirse también como algo causado por el querer encender la luz». Describir un suceso en términos de su causa no es confundir el suceso con su causa, ni tampoco la explicación en términos de una redescrición excluye la explicación causal.

El ejemplo sirve también para refutar la afirmación de que no podemos describir la acción sin usar palabras que la conecten con su supuesta causa. Aquí la acción ha de explicarse bajo la descripción: «mi presionar el interruptor», y la supuesta causa es «mi querer encender la luz». ¿Qué relación lógica pertinente se supone que existe entre estas dos expresiones? Parece más plausible defender un vínculo lógico entre «mi encender la luz» y «mi querer encender la luz», pero aun en este caso el vínculo resulta ser, tras una inspección, más bien gramatical que lógico.

De cualquier manera, hay algo muy extraño en la idea de que las relaciones causales son empíricas y no lógicas. ¿Qué puede querer decir esto? Seguramente no significa que todo enunciado causal verdadero sea empírico. Porque supóngase que «*A* causó *B*» es verdadero; entonces la causa de  $B = A$ ; por lo tanto, al sustituir resulta: «La causa de *B* causó *B*», lo cual es analítico. La verdad de un enunciado causal depende de *qué* sucesos se describan: su condición de analítico o sintético depende de *cómo* se describan los sucesos. No obstante, puede sostenerse que una razón racionaliza una acción sólo cuando se conforman apropiadamente las descripciones, y cuando las descripciones apropiadas no son lógicamente independientes.

Supongamos que decir que un hombre quiso encender la luz *significara* que él realizaría cualquier acción de la que creyera que lo llevaría a ese fin. Entonces el enunciado de su razón primaria para presionar el interruptor implicaría lógicamente que presionó el interruptor, que «actúa inmediatamente», como dice Aristóteles. En este caso, ciertamente habría una conexión lógica entre la razón y la

acción, el mismo tipo de conexión que la que hay entre «es soluble en agua y fue puesto en agua» y «se disolvió». Puesto que la implicación va de la descripción de la causa a la descripción del efecto, pero no a la inversa, nombrar la causa todavía proporciona información. Y aunque esta cuestión a menudo se pase por alto, «haberlo puesto en el agua causó su disolución» no implica lógicamente «es soluble en agua»; por tanto, esto último tiene una fuerza explicativa adicional. Sin embargo, la explicación sería mucho más interesante si, en lugar de referirnos a la solubilidad, con su conexión definicional obvia con el suceso que debe explicarse, pudiéramos referirnos a alguna propiedad, digamos a una estructura cristalina particular, cuya conexión con la disolución en el agua se conociera sólo mediante la experimentación. Ahora está claro por qué las razones primarias, como los deseos y anhelos, no explican las acciones de la manera relativamente trivial en que la solubilidad explica las disoluciones. La solubilidad, así lo suponemos, es una propiedad disposicional pura: se define en términos de un solo experimento. Pero no es posible definir los deseos por las acciones que puedan racionalizar, aun cuando la relación entre el deseo y la acción no sea simplemente empírica; hay otros criterios igualmente esenciales para los deseos: por ejemplo, su expresión en sentimientos y en acciones que no racionalizan. La persona que tiene un deseo, un anhelo o una creencia normalmente no necesita ningún criterio; generalmente sabe, aun en ausencia de toda clave accesible a los demás, lo que quiere, desea y cree. Estos rasgos lógicos de las razones primarias muestran que no es sólo una falta de ingenio lo que nos impide definir las como disposiciones para actuar por esas razones.

C. Según Hume, «podemos definir una causa como un objeto seguido de otro, y en donde todos los objetos similares al primero son seguidos por objetos similares al segundo». Pero Hart y Honoré sostienen que «el enunciado de que una persona hizo algo porque, por ejemplo, alguien la amenazó, no contiene ninguna implicación o afirmación implícita de que si las circunstancias se repitieran seguiría la misma acción» (p. 52). Hart y Honoré admiten que Hume tiene razón al decir que los enunciados causales singulares ordinarios implican generalizaciones, pero que se equivoca por esa misma razón, al suponer que los motivos y los deseos son causas ordina-

rias de las acciones. En suma, las explicaciones causales ordinarias esencialmente suponen leyes, pero no así las racionalizaciones.

Es común tratar de contestar este argumento sugiriendo que sí tenemos leyes rudimentarias que relacionan las razones con las acciones y que éstas en teoría pueden mejorarse. Ciertamente, las personas amenazadas no siempre responden de la misma manera; pero podemos distinguir entre amenazas y también entre agentes, en términos de sus creencias y actitudes.

No obstante, la sugerencia resulta engañosa porque las generalizaciones que conectan las razones con las acciones no son de la clase de leyes conforme a las cuales puedan hacerse confiadamente predicciones precisas, ni tampoco se pueden afinar y convertir en este tipo de leyes. Si reflexionamos sobre la manera en que las razones determinan la elección, la decisión y la conducta, es fácil ver por qué esto es así. Lo que surge como *la razón* en el aura *ex post facto* de la explicación y de la justificación frecuentemente era, para el agente en el momento de la acción, una consideración entre muchas, *una razón*. Toda teoría que sirva para predecir acciones sobre la base de razones debe encontrar una manera de evaluar la fuerza relativa de varios deseos y creencias en la matriz de la decisión; no puede tomar como su punto de partida el afirmamiento de lo que se espera de un solo deseo. El silogismo práctico agota su función al mostrar que una acción cae bajo una razón, por lo que no se puede afinar para convertirlo en una reconstrucción del razonamiento práctico que suponga la ponderación de razones que compiten entre sí. El silogismo práctico no ofrece un modelo ni para una ciencia predictiva de la acción, ni para una explicación normativa del razonamiento valorativo.

La ignorancia de leyes predictivas aptas no inhibe la explicación causal válida; de lo contrario podrían hacerse pocas explicaciones causales. Estoy seguro de que el cristal de la ventana se rompió porque una piedra lo golpeó, yo vi cómo sucedió todo; pero no dispongo (y ¿quién dispone?) de las leyes sobre cuya base pueda predecir qué golpes romperán qué ventanas. Una generalización como «los cristales de las ventanas son frágiles, y las cosas frágiles tienden a romperse cuando se las golpea con suficiente fuerza, en condiciones normales» no es una ley predictiva *grosso modo*; la ley predictiva, si la tuviéramos, sería cuantitativa y usaría conceptos

muy distintos. Esta generalización, como nuestras generalizaciones acerca de la conducta, sirve para una función diferente: ofrece pruebas en favor de la existencia de una ley causal que abarca el caso presente.<sup>7</sup>

Generalmente estamos mucho más seguros de una sola conexión causal que de cualquier ley causal que rija en ese caso; ¿demuestra esto que Hume estaba equivocado al afirmar que los enunciados causales singulares implican lógicamente leyes? No necesariamente, porque la tesis de Hume, tal como fue citada anteriormente, es ambigua. Puede significar que «A causó B» implica lógicamente alguna ley en particular que contiene los predicados usados en las descripciones «A» y «B», o puede significar que «A causó B» implica lógicamente que existe una ley causal ejemplificada por algunas descripciones verdaderas de «A» y «B».<sup>8</sup> Obviamente ambas versiones de la doctrina de Hume confieren un sentido a la afirmación de que los enunciados causales singulares implican lógicamente leyes y ambas sustentan la tesis según la cual las explicaciones causales «suponen leyes». Pero la segunda versión es mucho más débil en el sentido de que una afirmación causal singular no implica lógicamente ninguna ley en particular y puede defenderse, de ser necesario, sin defender ninguna ley. Se pueden ajustar la mayoría de las explicaciones causales sólo a la segunda versión de la doctrina de Hume; se adecua igualmente a las racionalizaciones.

La explicación más primitiva de un suceso ofrece su causa; las explicaciones más elaboradas podrán contar la historia más extensamente o podrán defender la afirmación causal singular al ofrecer una

7. En los ensayos 11, 12 y 13 se discuten temas que se mencionan en este párrafo y en el anterior.

8. Podríamos caracterizar a grandes rasgos el análisis de los enunciados causales singulares al que aludimos aquí de la siguiente manera: «A causó B» es verdadero si y sólo si hay descripciones de A y de B tales que la oración que se obtiene al poner estas descripciones en lugar de «A» y «B» en «A causó B» se sigue de una ley causal verdadera. Este análisis se salva de la trivialidad por el hecho de que no todas las generalizaciones verdaderas son leyes causales; las leyes causales se distinguen (aunque esto por supuesto no es un análisis) por el hecho de que se confirman inductivamente por sus ejemplificaciones y por el hecho de que apoyan enunciados causales singulares contrafácticos y subjuntivos. Hay más acerca de la causalidad en el ensayo 7.

ley pertinente o al dar razones para creer que ésta existe. Pero es un error pensar que no se habrá dado ninguna explicación mientras no se haya ofrecido una ley. Relacionada con estos errores está la idea de que los enunciados causales singulares apuntan necesariamente, por los conceptos que emplean, a los conceptos que aparecerán en la ley lógicamente implicada. Supongamos que un huracán del que se informa en la página 5 del *Times* del martes causa una catástrofe de la que se da información en la página 13 del *Tribune* del miércoles. Entonces el suceso relatado en la página 5 del *Times* del martes causó el suceso relatado en la página 13 del *Tribune* del miércoles. ¿Debemos buscar una ley que relacione sucesos de estas clases? Resulta apenas menos ridículo buscar una ley que relacione huracanes y catástrofes. A las leyes necesarias para predecir la catástrofe con precisión obviamente no les servirían conceptos como huracán y catástrofe. El problema que suscita la predicción del tiempo consiste en que las descripciones bajo las cuales nos interesan los sucesos —«un día frío, nublado, con una tarde lluviosa»— sólo tienen conexiones remotas con los conceptos empleados por leyes conocidas más precisas.

Las leyes cuya existencia se requiere si las razones son causas de las acciones no tratan con los mismos conceptos con los que deben tratar las racionalizaciones, de ello podemos estar seguros. Si las causas de una clase de sucesos (acciones) caen dentro de una clase determinada (razones) y si hay una ley para respaldar cada enunciado causal singular, de ello no se sigue que haya ley alguna que conecte los sucesos clasificados como razones con los sucesos clasificados como acciones; las clasificaciones pueden ser incluso neurológicas, químicas o físicas.

D. Se dice que el tipo de conocimiento que uno tiene de sus propias razones al actuar no es compatible con la existencia de una relación causal entre las razones y las acciones: una persona conoce de manera infalible sus propias intenciones al actuar, sin inducción ni observación, en tanto que ninguna relación causal ordinaria puede conocerse de este modo. Sin duda el conocimiento de nuestras propias intenciones al actuar mostrará muchas de las rarezas peculiares al conocimiento en primera persona que tenemos de nuestros propios dolores, creencias, deseos, etc.; el único problema es si estas rarezas



prueban que las razones no causan, en algún sentido ordinario, al menos, las acciones que racionalizan.

Puedes equivocarte fácilmente sobre la verdad de un enunciado de la forma «estoy envenenando a Carlos porque quiero evitarle dolor», porque puedes equivocarte acerca de si estás envenenando a Carlos —tú mismo puedes estar bebiendo la copa envenenada por error. Pero también parece que puedes equivocarte acerca de tus razones, particularmente cuando tienes dos razones para una acción, una que te gusta y otra que no. Por ejemplo, quieres evitarle un dolor a Carlos; también quieres deshacerte de él. Puedes equivocarte acerca de cuál fue el motivo que te hizo hacerlo.

El hecho de que puedas equivocarte no muestra que en general tenga sentido preguntarte cómo sabes cuáles fueron tus razones o que tenga sentido pedirte pruebas. Aunque puedas aceptar, en contadas ocasiones, pruebas públicas o privadas que muestren que te equivocas acerca de tus razones, por lo general ni tienes pruebas ni llevas a cabo ninguna observación. Por tanto, generalmente, el conocimiento de tus propias razones para actuar no es inductivo, porque cuando hay inducción, hay pruebas. ¿Muestra esto que este conocimiento no es causal? No veo cómo pueda mostrarlo.

Las leyes causales difieren de las generalizaciones verdaderas no legaliformes por el hecho de que sus ejemplificaciones las confirman; por tanto, la inducción es ciertamente un buen método para conocer la verdad de una ley. De esto no se sigue que sea la única manera. De cualquier modo, para saber que un enunciado causal singular es verdadero, no es necesario conocer la verdad de alguna ley; sólo es necesario saber que existe alguna ley que abarca los sucesos en cuestión. Y no es para nada evidente que la inducción y sólo la inducción produzca el conocimiento de que existe una ley causal que satisface ciertas condiciones; o para expresar lo mismo de otra manera, a menudo basta con un caso, tal como lo admitió Hume, para persuadimos de que existe una ley, y esto equivale a decir que nos persuadimos, sin pruebas inductivas directas, de la existencia de una relación causal.

*E.* Por último, me gustaría decir algo acerca de cierta incomodidad que algunos filósofos sienten al hablar de las causas de las acciones. Melden, por ejemplo, dice que a menudo las acciones

son idénticas a los movimientos corporales y que los movimientos corporales tienen causas; sin embargo, niega que las causas sean causas de las acciones. Me parece que esto es una contradicción. Lo lleva a ella el siguiente tipo de consideración: «Es inútil tratar de explicar la conducta por la eficacia causal del deseo: todo lo que *éste* puede explicar son acontecimientos ulteriores, pero no acciones realizadas por agentes. El agente que confronta el nexa causal en el que tales acontecimientos ocurren es una pobre víctima de todo lo que sucede en él y de todo lo que le ocurre a él» (pp. 128–129). Si no me equivoco, de ser válido, este argumento mostraría que las acciones no pueden tener causas en absoluto. No señalaré las dificultades obvias que genera el excluir completamente a las acciones del dominio de la causalidad. Pero quizá valga la pena tratar de poner al descubierto el origen del problema. ¿Por qué demonios una causa habría de convertir una acción en un mero acontecimiento y a una persona en una pobre víctima? ¿Es acaso porque tendemos a suponer, por lo menos en el campo de la acción, que una causa exige un causante y la acción un agente? Entonces insistimos en preguntar: si mi acción es causada, ¿qué la causó? Si fui yo, entonces se cae en el absurdo de la regresión infinita; si no fui yo, soy una víctima. Pero estas opciones obviamente no son exhaustivas. Algunas causas no tienen agentes. Entre estas causas sin agente están los estados y los cambios de estado en las personas, que por ser tanto razones como causas, convierten ciertos sucesos en acciones libres e intencionales.

## 2. ¿CÓMO ES POSIBLE LA DEBILIDAD DE LA VOLUNTAD?

La voluntad de un agente es débil si actúa, y actúa intencionalmente, en contra de su propio mejor juicio; en tales casos decimos, en ocasiones, que carece de fuerza de voluntad para hacer lo que sabe, o al menos cree, hechas todas las consideraciones, que sería lo mejor. Será conveniente denominar a las acciones de este tipo acciones incontinentes o decir que, al hacerlas, el agente actúa de manera incontinente. Al usar esta terminología me separo de la tradición, al menos al hacer más amplia de lo usual la clase de las acciones incontinentes. Pero la clase más amplia es la que deseo considerar y creo que incluye todas las acciones que algunos filósofos han denominado incontinentes y algunas de las acciones que muchos filósofos han denominado incontinentes.

Permítaseme explicar cómo mi concepción de la incontinencia es más general que algunas otras. Con frecuencia se hace que sea una condición de una acción incontinente el que se realice a pesar de que el agente sabe que otro curso de acción es mejor. Yo considero incontinentes tales acciones, pero la perplejidad que estudiaré depende sólo de la actitud o de la creencia del agente, así que insistir en el conocimiento restringiría el campo sin objeto. El conocimiento también tiene un innecesario y, por tanto, indeseado sabor de lo cognoscitivo; mi tema se refiere a juicios evaluativos, sea que éstos se analicen de manera cognoscitiva, prescriptiva o de otra manera. Así que incluso el concepto de creencia es, quizás, muy especial y yo hablaré de lo que el agente juzga o sostiene.

Si un hombre sostiene que, hechas todas las consideraciones, un curso de acción es el mejor, el correcto o la cosa que debe hacer y, sin embargo, hace alguna otra cosa, actúa de manera incontinente. Pero yo también diría que actúa de manera incontinente siempre que sostenga que es mejor, en general, algún curso disponible de acción diferente del que él sigue; o que juzga, entre un curso de acción que cree que le está disponible y la acción que realiza, que debería seguir el otro. En otras palabras, los juicios comparativos bastan para la incontinencia. Ahora podemos caracterizar una acción que revela debilidad de la voluntad o incontinencia:

D. Al hacer  $x$ , un agente actúa de manera incontinente si y sólo si: (a) el agente hace  $x$  intencionalmente; (b) el agente cree que hay una acción alternativa y que le es asequible; y (c) el agente juzga que, hechas todas las consideraciones, sería mejor hacer y que hacer  $x$ .<sup>1</sup>

Parece que hay acciones incontinentes en este sentido. La dificultad es su existencia pone en cuestión otra doctrina que tiene un aire de autoevidencia, a saber, que en tanto que una persona actúe intencionalmente actúa, como lo expresa Santo Tomás, a la luz de algún bien imaginado. Tal como se encuentra, esta tesis no contradice directamente la aseveración de que hay acciones incontinentes. Pero es difícil negar que las consideraciones que la favorecen, recomiendan también una versión relativizada: en la medida en que

1. En un útil artículo, G. Santas da esta versión de la incontinencia: «En un caso de debilidad un hombre hace algo que él sabe o cree que no habría que (debería) hacer, o no hace algo que sabe o cree que habría que hacer, cuando está presente la ocasión y la oportunidad para actuar o para abstenerse y cuando está en su poder, en algún sentido significativo, actuar conforme a su conocimiento o a su creencia» («Plato's *Protagoras* and Explanations of Weakness», p. 3). La mayoría de las diferencias entre esta descripción y la mía se deben a mi desviación deliberada de la tradición. Pero me parece que hay dos pequeños errores en la versión de Santas. Primero, la debilidad de la voluntad no requiere que la acción alternativa esté realmente disponible, sólo que el agente piense que lo está. Lo que busca Santas correctamente es que el agente actúe con libertad; pero, para esto, no es necesario que esté disponible para el agente la alternativa que éste piensa que es la mejor (o la que debe hacer). Por otra parte (y ésta es la segunda observación), los criterios de Santas no son suficientes para garantizar que el agente actúa intencionalmente, y pienso que esto es esencial a la incontinencia.

una persona actúe intencionalmente, actúa a la luz de lo que imagina (juzga) que es lo mejor.

Será útil explicitar esta tesis en la forma de dos principios. El primero expresa el supuesto natural acerca de la relación entre querer o desear algo y la acción. «El signo primitivo de querer es tratar de obtener», dice Anscombe en *Intention*.<sup>2</sup> Hampshire se aproxima más exactamente a lo que necesito cuando escribe, en *Freedom of the Individual*,<sup>3</sup> que «A quiere hacer *x*» es equivalente a «siendo iguales otras cosas, él haría *x*, si pudiera». Aquí tomo (posiblemente en contra de la intención de Hampshire) «siendo iguales otras cosas» como si significara, o de cualquier forma permitiera, la interpretación, «siempre que no haya algo que él quiera más». Dada esta interpretación, el principio de Hampshire podría, quizás, expresarse así:

P1. Si un agente quiere hacer *x* más que lo que él quiere hacer *y*, y se cree libre para hacer o bien *x* o bien *y*, entonces intencionalmente hará *x* si hace intencionalmente o bien *x* o bien *y*.

El segundo principio conecta juicios acerca de lo que es mejor hacer, con motivación o querer:

P2. Si un agente juzga que sería mejor hacer *x* que hacer *y*, entonces él quiere hacer *x* más que lo que él quiere hacer *y*.

De manera obvia, P1 y P2 implican juntos que si un agente juzga que sería mejor para él hacer *x* que hacer *y* y él mismo se cree libre para hacer o bien *x* o bien *y*, entonces intencionalmente hará *x* si hace *x* o *y* intencionalmente. Sugiero que esta conclusión parece mostrar que es falso que:

P3. Hay acciones incontinentes.

Alguien que esté convencido de que P1–P3 forman una tríada inconsistente, pero que encuentra realmente persuasivos sólo uno o dos de los principios, no tendrá dificultad en decidir qué decir. Pero para

2. G.E.M. Anscombe, *Intention*, p. 67. [Hay traducción al español: *Intención*, trad. Ana Isabel Stellino, Instituto de Investigaciones Filosóficas-UNAM/Paidós, México/Barcelona, 1991.]

3. S. Hampshire, *Freedom of the Individual*, p. 36.

alguien (como yo) a quien los principios P1–P3 le parecen evidentes de suyo, el problema que plantea la aparente contradicción es tan agudo como para llamarlo una paradoja. No puedo estar de acuerdo con Lemmon cuando escribe, en un artículo por lo demás admirable: «Quizás la *akrasia* es uno de los mejores ejemplos de un pseudo-problema en la literatura filosófica; en vista de su existencia, si lo ves como un problema ya habrás cometido un error filosófico.»<sup>4</sup> Si tus supuestos conducen a una contradicción, sin duda has cometido un error, pero puesto que puedes saber que has cometido un error sin saber cuál es el error, tu problema puede ser real.

Las soluciones propuestas que conozco al problema creado por la plausibilidad inicial de P1–P3 suponen que P1–P3 efectivamente se contradicen. Estas propuestas concluyen, naturalmente, abandonando uno u otro de los principios. No estoy muy contento con P1–P3 tal como los he formulado; quizás es fácil dudar si son verdaderos precisamente en esta forma (particularmente P1 y P2). Y reflexionando sobre las ambigüedades o la pluralidad de usos de varias de las palabras o frases críticas («juzgar mejor», «querer», «intencional») no es sorprendente que los filósofos hayan intentado interpretar alguna frase clave como si significara una cosa en un principio y significara algo distinto en otro. Pero yo estoy convencido de que por mucho que se remienden P1–P3, no se eliminará el problema subyacente; el problema sobrevivirá las nuevas formulaciones, reformamientos y eliminación de ambigüedades. Mencionaré algunas de las estrategias estándares e intentaré desacreditarlas, pero quedarán un sinfín de maneras de lidiar con el problema. Por tanto, mi estrategia básica no es la de intentar presentar un alegato inobjetable a favor de P1–P3, dándoles, quizás, una forma menos sujeta a excepciones. Más bien, lo que espero es mostrar que P1–P3 no se contradicen entre sí y, por tanto, que no debemos abandonar ninguno de ellos. Al mismo tiempo, ofreceré una explicación de por qué estamos inclinados a pensar que P1–P3 conducen a una contradicción; ya que, si tengo razón, un error común e importante explica nuestra confusión, un error acerca de la naturaleza de la razón práctica.

4. E.J. Lemmon, «Moral Dilemmas», pp. 144–145.

## I

He aquí algunas de las formas como los filósofos han intentado, o podrían intentar, enfrentarse al problema de la incontinencia, tal como lo he formulado. Los pecados del leopardo —lascivia, glotonería, avaricia y furia— son, conforme a Dante, los pecados menos serios por los que se nos puede condenar eternamente. Dante hace castigar estos pecados, que denomina los pecados de incontinencia, en los círculos segundo, tercero, cuarto y quinto del infierno. En un famoso ejemplo, Dante describe el pecado adúltero de Francesca da Rimini y Paolo Malatesta. Los comentaristas muestran su astucia señalando que incluso al narrar su historia Francesca muestra su debilidad de carácter. Así, Charles Williams dice: «Dante manipula la descripción, refuerza la excusa de tal manera, que la excusa se revela a sí misma precisamente como el pecado [...], la conversación persistente con la ocasión del pecado, la dulce y prolongada pereza del amor.»<sup>5</sup> Quizás todo esto sea verdadero de Francesca, pero no es esencial a la incontinencia, pues la «debilidad» puede ser momentánea, no un rasgo de carácter; cuando hablamos de «debilidad» podemos meramente expresar, sin explicarlo, el hecho de que el agente hizo lo que sabía que era erróneo («lo hizo un page»). Aristóteles incluso parece dar a entender que es imposible ser habitualmente incontinente, sobre la base de que la acción habitual comprende un principio conforme al cual uno actúa, mientras que el hombre incontinente actúa en contra de su principio. Supongo entonces que es al menos posible realizar acciones incontinentes aisladas y trataré la incontinencia como un hábito o vicio sólo en tanto que se interprete el vicio como el vicio de realizar con frecuencia, o habitualmente, acciones incontinentes.<sup>6</sup>

5. C. Williams, *The Figure of Beatrice*, p. 118.

6. «La incontinencia no es estrictamente un vicio [...] pues la incontinencia actúa en contra de la elección, el vicio conforme a ella» (*Et. Nic.* 1151a); «el vicio es como la hidropesía y la tuberculosis, en tanto que la incontinencia es como la epilepsia; el vicio es crónico, la incontinencia intermitente» (1150b). Pero *Donne* aparentemente describe el vicio de la incontinencia en uno de sus Sonetos Sagrados: «Oh to vex me, contraries meet in one; / Inconstancy unnaturally hath begot / A constant habit; that when I would not / I change in vows, and in devotion». [«Oh, para mi desazón, los contrarios se funden en uno; / De manera innatural, la

Un hombre podría sostener que es erróneo, hechas todas las consideraciones, enviarle un regalo de San Valentín a Estela Matutina. Empero, él puede enviarle un regalo de San Valentín a Estela Vespertina, y hacerlo intencionalmente, sin saber que Estela Vespertina era idéntica a Estela Matutina. Podríamos querer decir que él hizo algo que sostenía que era erróneo, pero sería engañoso decir que intencionalmente hizo algo que él sostenía que era erróneo; y el caso que ilustro ciertamente no es un ejemplo de una acción incontinente. Espero que sea claro que no debemos pensar que las acciones pueden agruparse, de manera simple, en incontinentes y otras. «Incontinente», como «intencional», «voluntario» y «deliberado», caracteriza acciones sólo en tanto que se las concibe de cierta manera más bien que de otra. En cualquier análisis serio de la forma lógica de las oraciones de acción, pienso que esas palabras deben interpretarse como operadores oracionales no funcionales veritativos: «Fue incontinente por parte de Francesca que. . . » y «Fue intencional por parte del agente que. . . » Pero, para los propósitos actuales, basta evitar el error de pasar por alto la intensionalidad de estas expresiones.

La incontinencia se caracteriza, con frecuencia, de una de las siguientes formas: el agente *tiene la intención de hacer y*, que él sostiene que es el mejor proceder o un mejor proceder que hacer  $x$ ;  $y$ , sin embargo, él hace  $x$ . O bien, el agente *decide hacer y*, que él sostiene que es el mejor proceder, o es un proceder mejor que hacer  $x$  y, sin embargo, él hace  $x$ . O bien, el agente *elige y* como resultado de la deliberación<sup>7</sup> y, sin embargo, hace  $x$ , que considera inferior a  $y$ . Cada una de estas formas de conducta es interesante y, dadas algunas especificaciones, puede caracterizarse como algo inconsistente, débil, vacilante o irracional. Cualquiera de ellas podría ser un caso de incontinencia, tal como la he definido. Pero, tal como se

inconstancia ha engendrado / Un hábito constante: que, cuando yo no debiera, / Cambio de compromisos y de devoción»].

7. En ocasiones, Aristóteles caracteriza al hombre incontinente (al *ákrates*) como alguien que «abandona su elección» (*Et. Nic.*, 1151a) o como alguien que «abandona la conclusión a la que ha llegado» (1145b); pero también con frecuencia, conforme a la línea que aquí se sugiere: «hace lo que sabe que es malo» (1134b) o «está convencido de que debe hacer una cosa y, sin embargo, hace otra cosa» (1146b).



encuentran, no son necesariamente casos de incontinencia porque ninguna de ellas implica formalmente que cuando el agente actúa, éste sostiene que, hechas todas las consideraciones, otro curso de acción sería mejor. Y, por otra parte, una acción puede ser incontinente sin que el agente jamás haya decidido, elegido o pretendido hacer lo que él juzga que es mejor.

El principio 2 enuncia una forma suave de internalismo. Dice que un juicio de valor debe reflejarse en lo que el agente quiere (o en sus deseos o motivos). Esto no es tan fuerte como muchas formas de internalismo; por ejemplo, no dice nada en absoluto acerca de la conexión entre el valor real de las cosas (o el carácter obligatorio de las acciones) y los deseos o motivos. Ni, hasta donde puedo ver, nos compromete con una doctrina acerca de lo que significan los juicios evaluativos. Conforme a Hare, «llamar la atención hacia las íntimas relaciones lógicas entre, por una parte, querer algo y pensar que es bueno y, por otra, entre querer y hacer algo con relación a obtener lo que uno quiere, es caer en manos del prescriptivista, pues es añadir un nuevo eslabón entre considerar que algo es bueno y la acción».<sup>8</sup> Confieso que no veo cómo estas «íntimas relaciones lógicas», que P1 y P2 proporcionan de cierta manera, apoyan cualquier teoría particular acerca del significado de oraciones o términos evaluativos. Una fuente posible de confusión se revela cuando Hare dice: «si los juicios morales no fuesen prescriptivos, no habría problema alguno acerca de la debilidad moral; pero hay un problema; por tanto, son prescriptivos» (p. 68). La confusión se da entre hacer un juicio y el contenido del juicio. Es P2 (o sus similares) el que crea el problema y P2 conecta *hacer* un juicio con querer y, de esto, *vía* P1, con actuar. Pero el prescriptivismo es una doctrina acerca del *contenido* o *significado* de lo que se juzga y P2 nada dice acerca de esto. Por ejemplo, uno podría sostener que decir que un curso de acción es mejor que otro es decir tan sólo que creará más placer, y sostener, sin embargo, como Mill quizás lo hizo, que quienquiera que crea que cierto curso de acción creará más placer que otro (necesariamente) lo quiere más. Así, a mí me agradaría negar que hay una conexión simple entre el problema de la incontinencia, tal como lo he formulado, y cualquier teoría ética particular.

8. R.M. Hare, *Freedom and Reason*, p. 71.

Quizás la manera más común de tratar el problema de la incontinencia es rechazar P2. Después de todo, parece bastante obvio que podemos pensar que *x* es mejor *y*, sin embargo, querer más *y*. Incluso P2 es más fácil de poner en cuestión si se enuncia en la forma: si un agente piensa que debe (o que está obligado a) hacer *x*, entonces él quiere hacer *x*, pues, claro está, con frecuencia no queremos hacer lo que pensamos que debemos hacer. Hare, si es que lo entiendo, da cuenta de algunos casos de incontinencia de esa manera; de acuerdo con Santas, así lo hizo Platón.<sup>9</sup>

Es fácil interpretar P2 de tal manera que resulte falso, pero es más difícil creer que no hay una lectura natural que lo haga verdadero, pues, en contra de nuestra tendencia a aceptar que con frecuencia creemos que debemos hacer algo *y*, sin embargo, no queremos hacerlo, también está la tendencia opuesta a decir que si alguien cree realmente (sinceramente) que debe, entonces su creencia debe mostrarse en su conducta (y por esto, claro está, en su inclinación o en su deseo a actuar). Cuando nos importa contrastar pensar que debemos con querer, continúa esta línea de pensamiento, o bien estamos usando la frase «pensar que debemos» con el significado de algo similar a «pensar que es lo que se requiere conforme a los patrones usuales de la comunidad», o estamos restringiendo el querer a lo que es atractivo sobre una base puramente egoísta o personal. Aun cuando encuentro atractivas estas formas de defender P2, las mismas son difíciles de hacer concluyentes sin cometer una petición de principio en contra de la cuestión presente. Así que estoy inclinado, a fin de avanzar, a señalar que de alguna forma surgirá un problema acerca de la incontinencia mientras haya alguna palabra o frase que convincentemente podamos sustituir por «quiere», tanto en P1 como en P2.

Otra línea común de pensamiento que se puede tomar con la incontinencia es describir al *ákrates* como dominado por la pasión o trastornado por la emoción. «Ciertamente sé qué mal pretendo hacer. Pero más fuerte es mi furia que todas mis cavilaciones», vocifera Medea. Hare hace de éste el paradigma de todos los casos de debilidad de la voluntad, donde no podemos simplemente separar el juicio moral y el deseo, y añade que, en tales casos, el agente es

9. G. Santas, «The Socratic Paradoxes».

psicológicamente incapaz de hacer lo que piensa que debe (*Freedom and Reason*, p. 77). Hare cita la *Medea* de Eurípides cuando ella dice: «una compulsión desconocida, muy a mi pesar, me acosa» y a San Pablo cuando escribe: «el bien que quiero hacer no lo hago, sino que lo que hago es el mal que va en contra de mi voluntad; y si lo que hago va en contra de mi voluntad, claramente ya no soy el agente» (*Romanos 7*). Esta línea de pensamiento conduce a la tesis de que uno nunca actúa intencionalmente en contra del propio mejor juicio y, así, niega P3; no hay acciones incontinentes en el sentido que hemos definido.<sup>10</sup>

Una tesis relacionada con la anterior, pero diferente, es la de Aristóteles, según la cual la pasión, la lascivia o el placer distorsionan el juicio y, así, impiden que un agente se forme un juicio pleno de que su acción es errónea. Aun cuando hay mucho lugar para la duda acerca de cuál fue precisamente la tesis de Aristóteles, es seguro decir que intentó resolver nuestro problema distinguiendo dos sentidos en los que puede decirse que un hombre sabe (o cree) que una cosa es mejor que otra; un sentido hace verdadero a P2, en tanto que el otro sentido se necesita en la definición de incontinencia. El sabor de este segundo sentido lo da la observación de Aristóteles de que el hombre incontinente tiene conocimiento «en el sentido en el que tener conocimiento no significa saber, sino sólo hablar, como un borracho puede musitar los versos de Empédocles» (*Et. Nic.* 1147b).

Quizás es evidente que hay una considerable gama de acciones similares a las acciones incontinentes en uno u otro respecto, donde podemos hablar de autoengaño, insinceridad, mala fe, hipocresía, deseos, motivos e intenciones inconscientes, etc.<sup>11</sup> De hecho, hay una tentación muy grande, al trabajar en este tema, de jugar al

10. Santo Tomás es excelente acerca de esto. Él distingue claramente entre acciones realizadas debido a una fuerte emoción, tales como el temor, que concede que, en cierta medida, son involuntarias y, por tanto, no verdaderamente incontinentes, y acciones que se realizan por concupiscencia, por ejemplo; aquí, él dice, «la concupiscencia inclina la voluntad a desear el objeto de la concupiscencia. Por tanto, el efecto de la concupiscencia es hacer que algo sea voluntario» (*Summa Theologica*, Parte II, Q. 6).

11. «No es sino una prisa banal la que concluye insinceridad de lo que observadores externos llaman inconsistencia.» (George Eliot, *Middlemarch*).

psicólogo aficionado. Nos estamos muriendo por decir: recuerda la enorme variedad de formas en las que un hombre puede creer o sostener algo, saberlo, querer algo, tenerle miedo o hacer algo. Podemos actuar como si supiésemos algo y, sin embargo, dudarlo profundamente; podemos actuar al límite de nuestra capacidad y, al mismo tiempo, separarnos como observadores y decimos a nosotros mismos: «qué extraño es hacer eso». Podemos desear cosas y decimos a nosotros mismos que las odiamos. Estos estados a medias y estados contradictorios son comunes y plenos de interés para el filósofo. Sin duda explican, o al menos señalan un camino para describir sin contradicción muchos casos en los que nos encontramos hablando de debilidad de la voluntad o de incontinencia. Pero nosotros mismos mostramos cierta debilidad como filósofos si no llegamos a preguntar: ¿comprende cada caso de incontinencia una de estas zonas sombrías donde deseamos tanto aplicar como suspender el uso de algún predicado mental? ¿Nunca sucede que yo tenga un juicio no empañado, no vacilante de que mi acción, hechas todas las consideraciones, no es lo mejor y, sin embargo, que la acción que yo realizo no tenga rastro alguno de carácter compulsivo o coercitivo? No hay prueba de que existan tales acciones; pero me parece absolutamente cierto que existen. Y, si esto es así, por más atención que prestemos a los sutiles casos límite de porciones de conducta, no se resolverá el problema central.<sup>12</sup>

Austin se queja de que al tratar el tema presente estamos propensos a «sufrir un colapso, sucumbiendo a la tentación de perder el control de nosotros mismos». El autor amplía lo anterior:

Supongo que Platón, y tras él Aristóteles, nos impuso esta confusión, tan mala en su momento y a su manera, como la última, grotesca, confusión de debilidad moral con debilidad de la voluntad.

12. «Oh, dime, ¿quién lo declaró por primera vez, quién proclamó por vez primera que el hombre sólo hace cosas repulsivas porque no conoce sus propios intereses reales [...]? ¿Qué hay que hacer con los millones que dan testimonio de que los hombres, a sabiendas, esto es, entendiendo plenamente sus intereses reales, les han dado la espalda y se han lanzado precipitadamente por otro camino [...] impulsados a seguir este camino por nadie y por nada?» (Dostoievski, *Notas del subsuelo*).

Me atraen mucho los helados y en la mesa principal <sup>13</sup> se sirve un postre helado dividido exactamente en segmentos para cada uno de los comensales; yo estoy tentado a servirme dos segmentos y lo hago, sucumbiendo así a la tentación e incluso, concebiblemente (pero ¿por qué necesariamente?), yendo en contra de mis principios. Pero ¿pierdo el control de mí mismo? ¿Devoro, cojo los trozos del platón y me lo trago, ajeno a la consternación de mis colegas? Nada de eso. Con frecuencia sucumbimos a la tentación con calma e incluso con delicadeza.<sup>14</sup>

Sucumbimos a la tentación con calma; también hay muchos casos en los que actuamos en contra de nuestro mejor juicio y que no pueden describirse como sucumbir a la tentación.

Comienza a aclararse que en las versiones usuales de la incontinencia hay dos temas muy distintos que se entrecruzan y tienden a confundirse. Uno es que el deseo nos distrae del bien o nos obliga al mal; el otro es que la acción incontinente siempre favorece la pasión bestial, egoísta, sobre el llamado del deber y de la moralidad. Que estos dos temas pueden separarse lo subrayó Platón, tanto en el *Protágoras* como en el *Filebo*, cuando mostró que el hedonista, tan sólo por su propia inclinación hacia el placer, podía ir en contra de su propio mejor juicio tan fácilmente como cualquier otro. Mill señala lo mismo, aun cuando, presumiblemente, desde una posición más favorable al hedonista: «Con frecuencia los hombres, por flaqueza de carácter, eligen el bien más cercano, aun cuando saben que es el menos valioso; y esto sucede por igual cuando la elección es entre dos placeres corpóreos que cuando es entre lo corpóreo y lo mental» (*Utilitarismo*, Cap. 11). Desafortunadamente, Mill echa a perder el efecto de esta propuesta añadiendo: «Ellos persiguen las satisfacciones sensuales en detrimento de la salud, aun cuando son perfectamente conscientes de que la salud es el bien mayor.»

Como un primer paso positivo para tratar el problema de la incontinencia, propongo divorciar completamente ese problema de la preocupación del moralista de que un placer vívido puede acallar, embotar o engañar nuestro sentido de lo convencionalmente correc-

13. *High Table* en la Universidad de Oxford. [N. del T.]

14. J.L. Austin, «A Plea for Excuses», p. 146.

to. Acabo de tenderme relajadamente en la cama, tras un día difícil, cuando se me ocurre que no me he lavado los dientes. La preocupación por mi salud me pide que me levante y me lave; la satisfacción sensual sugiere que olvide los dientes por una vez. Peso las alternativas a la luz de las razones: por una parte, mis dientes están fuertes y, a mi edad, el decaimiento es lento. No importaría mucho si no me los cepillara. Por otra parte, si me levanto, esto podría estropear mi calma y podría tener como resultado una mala noche de sueño. Considerándolo todo, juzgo que yo haría mejor quedándome en cama. Sin embargo, mi sentimiento de que debo cepillarme los dientes es demasiado fuerte para mí; con fastidio dejo mi cama y cepillo mis dientes. Mi acto es claramente intencional, aun cuando va en contra de mi mejor juicio y es por ello incontinente.

Hay numerosas ocasiones en las que el placer inmediato cede ante los principios, la cortesía o el sentido del deber y, sin embargo, juzgamos (o sabemos) que, hechas todas las consideraciones, deberíamos optar por el placer. Al abordar el problema de la incontinencia, una buena estrategia es detenemos ante los casos en los que la moralidad simplemente no figura en el cuadro como uno de los contendientes que solicitan nuestra atención —o, si lo hace, se halla del lado equivocado. Entonces no sucumbiremos a la tentación de reducir la incontinencia a casos tan especiales tales como el ser dominados por la bestia dentro de nosotros, no escuchar el llamado del deber o caer en la tentación.<sup>15</sup>

15. No conozco ningún caso claro de un filósofo que reconozca que la incontinencia no es esencialmente un problema de filosofía moral, sino un problema de la filosofía de la acción. Butler, en los *Sermons* (párrafo 39, prefacio a «The Fifteen Sermons Preached at the Rolls Chapel»), señala que «la benevolencia hacia personas particulares puede ser un grado de debilidad y, por eso, reprehensible», pero aquí suena muy fuerte la nota de autocomplacencia. Y Nowell-Smith, *Ethics*, pp. 243 ss., describe muchos casos de incontinencia en los que nos dominan la conciencia o el deber: «Paradójicamente, pero no injustamente, podríamos decir que en tal caso es difícil resistir la tentación de decir la verdad. Somos los esclavos de nuestras propias conciencias.» Los esclavos no actúan libremente; nuevamente, el caso no es claro.

Aristóteles trata el caso del hombre que, en contra de su propio principio (y de su mejor juicio) persigue (vehementemente) algo noble y bueno (le preocupa mucho el honor o sus hijos), pero se niega a llamar esto incontinencia (*Et. Nic.*, 1148).

## II

Bajo el escrutinio escéptico, P1 y P2 aparecen muy vulnerables y, sin embargo, manipularlos no produce ninguna versión satisfactoria acerca de cómo es posible la incontinencia. Al menos parte de la razón se halla en el hecho de que P1 y P2 derivan su fuerza de una tesis muy persuasiva acerca de la naturaleza de la acción intencional y del razonamiento práctico. Cuando una persona actúa con una intención, la siguiente parece ser una descripción verdadera, aunque aproximada e incompleta, de lo que pasa: ella le da un valor positivo a algún estado de cosas (un fin o la realización, por ella misma, de una acción que satisfaga ciertas condiciones); ella cree (sabe o percibe) que una acción, de un tipo que ella puede hacer, promoverá, producirá o realizará el estado de cosas que valora y, así, ella actúa (esto es, actúa *debido* a su valoración o deseo y a su creencia). Esta descripción, generalizada y refinada, les ha parecido a muchos filósofos, de Aristóteles en adelante, que promete dar un análisis de lo que es actuar con una intención; que ilumina cómo explicamos una acción dando las razones que tuvo el agente al actuar y que proporciona el principio de una explicación del razonamiento práctico, esto es, del razonamiento acerca de qué hacer, del razonamiento que conduce a la acción.

En el caso más simple, imaginamos que el agente tiene, por ejemplo, un deseo de saber la hora. Él se da cuenta de que puede satisfacer su deseo mirando su reloj, así que mira su reloj. Podemos responder la pregunta de por qué miró su reloj; sabemos la intención con la que lo hizo. Siguiendo a Aristóteles, el deseo puede concebirse como un principio de acción y su expresión proposicional natural sería aquí algo como «sería bueno que yo supiese la hora» o incluso, de manera más forzada «es deseable cualquier acto mío que tenga como resultado el que yo sepa la hora». Aristóteles compara un principio así con la premisa mayor de un silogismo. La expresión proposicional de la creencia del agente sería, en este caso, «mirar mi reloj tendrá como resultado el que yo sepa la hora»; esto corresponde a la premisa menor. Subsumiendo el caso bajo la regla, el agente realiza la acción deseable: mira su reloj.

Parece que, dados este deseo y esta creencia, el agente está en posición de inferir que es deseable ver su reloj y, de hecho, hacer

tal inferencia es algo que sería natural describir como subsumir el caso bajo la regla. Pero, dados el deseo y esta creencia, también se satisfacen condiciones que conducen a (y por tanto explican) una acción intencional, y así Aristóteles dice que una vez que una persona tiene el deseo y cree que alguna acción lo satisfará, *actúa de inmediato*. Puesto que no hay forma de distinguir las condiciones bajo las que un agente está en posición de inferir que es deseable una acción que él está en libertad de realizar, de las condiciones bajo las que actúa, Aristóteles aparentemente identifica el extraer la inferencia y el actuar; nos dice: «la conclusión es una acción». Pero está claro que esta versión de la acción intencional y de la razón práctica contradice el supuesto de que hay acciones incontinentes.

Mientras tengamos a la vista el esbozo general de la teoría de Aristóteles, pienso que no podemos dejar de darnos cuenta de que no puede ofrecer ningún análisis satisfactorio de la acción incontinente. Sin duda puede explicar por qué, en casos límite, estamos tentados a decir tanto que un agente actuó intencionalmente como que sabía que no debía. Pero si postulamos un deseo fuerte con base en el cual actuó, entonces, conforme a la teoría, también le atribuimos al agente un juicio fuerte de que la acción es deseable; y si subrayamos que la habilidad del agente para razonar acerca de lo incorrecto de su acción se debilitó o distorsionó, en esa medida mostramos que no apreció plenamente que lo que estaba haciendo era indeseable.

No ha de suponerse que podemos escapar a la dificultad de Aristóteles tan sólo abandonando la doctrina de que tener razones para actuar siempre tiene como resultado la acción. Por ejemplo, podríamos conceder que una persona puede tener un deseo y creer que una acción lo satisfará y, sin embargo, no llega a actuar, y añadir que sólo si el deseo y la creencia causan su acción podemos hablar de una acción intencional.<sup>16</sup> Conforme a esta versión modificada de la teoría de Aristóteles (si es realmente una modificación), aún tendríamos que explicar por qué en algunos casos el deseo y la creencia causan una acción, mientras que en otros tan sólo conducen al juicio de que era deseable cierto curso de acción.

El hombre incontinente cree que, en general, sería mejor hacer alguna otra cosa, pero él tiene una razón para lo que hace, pues su

16. Para una versión de esta teoría véase el ensayo 1.



acción e intencional. Por tanto, debemos ser capaces de abstraer de su conducta y de su estado mental un razonamiento práctico cuya conclusión es, o sería si la conclusión se extrajese de las premisas, que es deseable la acción de hecho realizada. Aristóteles tiende a oscurecer este aspecto concentrándose en casos en los que el hombre incontinente se comporta «bajo la influencia de la regla y de la opinión» (*Et. Nic.*, 1147b; *cfr.* 1102b).

Santo Tomás es mucho más claro que Aristóteles en este importante asunto. Nos dice:

Quien tiene conocimiento del universal está impedido, por una pasión, de razonar a la luz de ese universal, de tal manera que extraiga la conclusión; pero razona a la luz de otra proposición universal sugerida por la inclinación de la pasión y extrae su conclusión conforme a esto [...] Por tanto, la pasión encadena la voluntad y le impide que piense y concluya conforme a la primera proposición, de tal manera que mientras la pasión dure, la razón arguye y concluye conforme a la segunda.<sup>17</sup>

Un ejemplo de Santo Tomás muestra los apuros del hombre incontinente:

EL LADO DE LA RAZÓN	EL LADO DE LA LUJURIA
(M <sub>1</sub> ) Ninguna fornicación es lícita	(M <sub>2</sub> ) Ha de perseguirse el placer
(m <sub>1</sub> ) Éste es un acto de fornicación	(m <sub>2</sub> ) Éste es un acto placentero
(C <sub>1</sub> ) Este acto no es lícito	(C <sub>2</sub> ) Este acto ha de perseguirse

Podemos presentar esto de manera más perturbadora, aun cuando al hacerlo vayamos más allá de Aristóteles y de Santo Tomás, si interpretamos los principios y las conclusiones como juicios comparativos acerca de los méritos de cometer o no cometer el acto en cuestión. Las dos conclusiones, (C<sub>1</sub>) y (C<sub>2</sub>), serán entonces (dados

17. *Summa Theologica*, Parte II, Q. 77, Art. 2, réplica a la objeción 4. Santo Tomás cita al apóstol: «Veo otra ley en mis miembros en lucha en contra de la ley de mi mente.»

algunos supuestos naturales): es mejor no realizar este acto que realizarlo, y: es mejor realizar este acto que no realizarlo. Y éstas se encuentran en contradicción flagrante bajo el supuesto de que «mejor que» es asimétrico.

Y ahora debemos observar que *esta imagen del razonamiento moral no es tan sólo inadecuada para dar una versión de la incontinencia; no puede dar una versión correcta de los casos simples de conflicto moral*. Por un caso de conflicto moral me refiero a un caso en el que hay buenas razones tanto para realizar una acción como para realizar una que la excluya (quizás absteniéndose de la acción). Hay conflicto en este sentido mínimo siempre que el agente sea consciente de consideraciones que, por sí solas, conducirían a acciones mutuamente incompatibles; sentimientos de lucha y de ansiedad son maquillaje inesencial. Es bastante claro que la incontinencia puede existir sólo cuando hay conflicto en este sentido, pues el hombre incontinente sostiene que un curso es el mejor (por una razón) y, sin embargo, hace alguna otra cosa (también por una razón). Así que podemos dejar de lado, por el momento, lo que le es específico a la incontinencia y considerar el conflicto en general. Los argumentos mellizos del párrafo anterior describen no sólo el apuro del hombre incontinente, sino también el del hombre recto en lucha con la tentación; uno de ellos hace la cosa equivocada y el otro la correcta, pero ambos actúan ante demandas en conflicto.

La situación es común; la vida está llena de ejemplos: debo hacerlo porque salvaré una vida, no debo hacerlo porque será una mentira; si lo hago, no mantendré la palabra que le di a Lavinia, si no lo hago, no mantendré la palabra que le di a Lolita; y así sucesivamente. Cualquiera puede encontrarse en este predicamento, por más que sea honesto o contemporizador, de voluntad débil o fuerte. Pero entonces, a menos que adoptemos la línea de que los principios morales no pueden entrar en conflicto en la aplicación a un caso, debemos abandonar el concepto de la naturaleza de la razón práctica que hasta ahora hemos estado asumiendo, pues ¿cómo es que las premisas, todas las cuales son verdaderas (o aceptables) implican formalmente una contradicción?

Es sorprendente que en la filosofía moral contemporánea este problema haya recibido poca atención y ningún tratamiento satisfactorio. Quienes reconocen la dificultad parecen estar dispuestos

a aceptar una de dos soluciones: conceden, en efecto, un único principio moral último, o bien se contentan con la noción de una distinción entre lo deseable *prima facie* (el bien, lo obligatorio, etc.) y lo absolutamente deseable (el bien, lo obligatorio, etc.).<sup>18</sup> No argumentaré aquí la propuesta, pero no creo que pueda aceptarse ninguna versión de la solución de «principio único» una vez que se entiendan sus implicaciones: los principios o las razones para actuar son irreductiblemente múltiples. Por otra parte, no es fácil ver cómo sacar ventaja de la pretendida distinción entre valores *prima facie* y absolutos. Supongamos, primero, que intentamos pensar en «*prima facie*» como un adverbio atributivo, que ayuda a formar predicados tales como «*x es prima facie* bueno, correcto, obligatorio» o «*x es, prima facie, mejor que y*». Para evitar nuestro problema reciente, debemos suponer que «*x es, prima facie, mejor que y*» no contradice a «*y es, prima facie, mejor que x*» y que «*x es prima facie* correcto» no contradice a «*x es prima facie* incorrecto». Pero, entonces, la conclusión que podemos extraer en cualquier clase de conflicto (y, por tanto, de incontinencia) será «*x es, prima facie, mejor que y y y es, prima facie, mejor que x*». Esto se reduce a decir, como lo hace claro la estructura que el razonamiento práctico tendría conforme a este supuesto, «hay algo que decir a favor y hay algo que decir en contra de hacer tal y cual —y también a favor y en contra de no hacerlo». Probablemente esto pueda decirse acerca de cualquier acción que sea; de cualquier manera, es difícil aceptar la idea de que la suma de nuestra sabiduría moral acerca de qué hacer en una situación dada tiene esta forma. La situación que describo no se altera de ninguna manera interesante si «*prima facie*» u «obligatorio *prima*

18. Ejemplos de tesis que, en efecto, conceden sólo un principio moral último: Kurt Baier, en *The Moral Point of View*, sostiene que en casos de conflicto entre principios hay principios de orden superior que señalan qué principios tienen precedencia; Singer, «Moral Rules and Principles», sostiene que los principios morales no pueden entrar en conflicto; Hare, en *The Language of Morals*, arguye que no hay excepciones a principios morales aceptables. Si los principios últimos nunca entran en conflicto o nunca tienen contraejemplos, podemos aceptar la conjunción de principios últimos como nuestro principio único, mientras que si hay un principio de orden superior que resuelve los conflictos, obviamente podemos construir un principio sin excepciones, único. Y, claro está, todos los utilitaristas francos, sean o no de regla, creen que hay un principio moral, único, sin excepciones.

*facie*» se trata como un operador oracional (no funcional veritativo) más bien que como un predicado. En breve volveré a este problema; reconsideremos ahora la incontinencia.

La imagen de la incontinencia que obtenemos de Aristóteles, Santo Tomás y Hare es la de una batalla o lucha entre dos contendientes. Cada contendiente está armado con su argumento o principio. Uno de los lados puede denominarse «pasión» y el otro «razón»; ambos luchan entre sí; un lado gana, el lado equivocado, el lado denominado «pasión» («lujuria» o «placer»). Empero, hay otra imagen en competencia (que puede encontrarse en Platón, así como en Butler y en muchos otros). Quizás la presagia Dante (quien cree seguir a Santo Tomás y a Aristóteles) cuando habla del hombre incontinente como de quien «deja que el deseo saque a la razón de su trono» (*Inferno*, Canto v). Aquí hay tres actores en escena: la razón, el deseo y quien deja que el deseo obtenga la ventaja. El tercer actor quizás se denomine «la Voluntad» (o «Conciencia»). Depende de la Voluntad decidir quién gana la batalla. Si la Voluntad es fuerte, le da la palma a la razón; si es débil, puede darle el triunfo al placer o a la pasión.

Sugiero que esta segunda imagen es superior a la primera, por absurdas que ambas nos puedan parecer. Conforme a la primera historia no sólo no podemos dar cuenta de la incontinencia; no es claro cómo podemos jamás culpar al agente por lo que hace: su acción meramente refleja el resultado de una lucha interior. ¿Qué más puede hacer? Y, más importante, la primera imagen no nos permite darle sentido a un conflicto en el alma de una persona, porque no deja espacio para el muy importante proceso de sopesar las consideraciones.<sup>19</sup> Conforme a la segunda imagen, la representante del agente, la Voluntad, puede juzgar la fuerza de los argumentos de ambos lados, puede ejecutar la decisión y recibir el castigo. El único problema es que parece que volvemos a donde empezamos, pues ¿cómo puede la Voluntad juzgar que una vía de acción es la mejor y, sin embargo, escoger la otra?

Sería un error pensar que no hemos progresado nada. Pues lo que ahora deben sugerir estas coloridas metáforas de gladiadores y

19. Una versión más compleja del conflicto, que parece hacer surgir el mismo problema, es la versión de Ryle en *The Concept of Mind*, pp. 93–95.

jueces es que, en el conflicto moral y, por tanto, en la incontinencia, está presente un razonamiento práctico que, hasta ahora, hemos pasado completamente por alto. Lo que debe añadirse a la imagen es un nuevo argumento:

LA VOLUNTAD (CONCIENCIA)

( $M_3$ )  $M_1$  y  $M_2$

( $m_3$ )  $m_1$  y  $m_2$

( $C_3$ ) Esta acción es equivocada

Es claro que resulta necesario algo como este tercer argumento si un agente ha de actuar, sea correcta o incontinentemente, frente a un conflicto. No basta saber las razones de cada lado: él debe saber a qué equivalen.<sup>20</sup> El hombre incontinente va en contra de su mejor juicio y esto, seguramente, es ( $C_3$ ), que se basa en todas las consideraciones, y no ( $C_1$ ), que no logra tomaren cuenta las razones del otro lado. Se podría decir que hemos descubierto dos significados muy diferentes de la frase «su mejor juicio». Esto podría significar cualquier juicio a favor del lado correcto (razón, moralidad, familia, país); o el juicio basado en todas las consideraciones pertinentes que el actor conoce. He argumentado que la primera noción no es realmente pertinente para el análisis de la incontinencia.

Pero ahora nos enfrentamos a nuestro otro problema, la forma o la naturaleza del razonamiento práctico. Pues nada podría ser más obvio que el que nuestro tercer «silogismo práctico» no es silogismo alguno; la conclusión simplemente no se sigue lógicamente de las premisas. Y, de todas maneras, introducir ese tercer razonamiento no resuelve el problema que teníamos antes: aún tenemos «conclusiones» contradictorias. En este punto podríamos intentar introducir, una vez más, «*prima facie*» en lugares adecuados: por ejemplo, en ( $M_1$ ), ( $M_2$ ), ( $C_1$ ) y ( $C_2$ ). Luego podríamos intentar relacionar la deseabilidad *prima facie* con la deseabilidad sin más al hacer que ( $C_1$ ) y ( $C_2$ ), así interpretados, fuesen los datos para ( $C_3$ ). Pero ésta no es una línea de pensamiento promisoria. Difícilmente podemos

20. Mi autoridad acerca de a qué equivalen en este caso es Santo Tomás; véase la referencia en la nota 17.

llegar a saber si una acción debe realizarse simplemente por el hecho de que es tanto *prima facie* correcta como *prima facie* incorrecta.

Ahora es clara la fuente real de la dificultad: si hemos de tener una teoría coherente de la razón práctica, debemos abandonar la idea de que podemos *separar* las conclusiones acerca de lo que es deseable (o mejor) u obligatorio de los principios que les dan color a esas conclusiones. El problema yace en el supuesto tácito de que los principios morales tienen la forma de condicionales universalizados; una vez que se hace este supuesto, nada que podamos hacer con un operador *prima facie* en la conclusión salva las cosas. A este respecto, la situación es como razonar a partir de evidencia probabilista. Como lo ha subrayado Hempel con gran claridad,<sup>21</sup> no podemos razonar a partir de:

( $M_4$ ) Si el barómetro desciende, casi de seguro lloverá

( $m_4$ ) El barómetro está descendiendo

a la conclusión:

( $C_4$ ) Casi de seguro lloverá

porque podemos, al mismo tiempo, estar igualmente justificados para argumentar:

( $M_5$ ) Cielos rojos en la noche, casi de seguro no lloverá

( $m_5$ ) El cielo está rojo esta noche

∴ ( $C_5$ ) Casi de seguro no lloverá

La falla crucial es interpretar ( $M_4$ ) y ( $M_5$ ) de tal manera que permitan la separación de la conclusión modal. Una forma de arreglar el asunto es considerar que el «casi de seguro» de ( $M_4$ ) y ( $M_5$ ) no modifica la conclusión sino la conectiva. Así, podríamos expresar ( $M_4$ ) como «que el barómetro desciende hace probable que llueva»; en símbolos, «*pr* ( $Lx, Dx$ )», donde la variable tiene como rango áreas de espacio-tiempo que pueden estar caracterizadas por barómetros en descenso o por lluvia. Si hacemos que «*a*» nombre el espacio y el tiempo de aquí y ahora y que « $Cx$ » signifique que la parte temprana de  $x$  está caracterizada por un cielo rojo en la tarde,

21. Carl Hempel, *Aspects of Scientific Explanation*, pp. 394–403.

podemos intentar reconstruir de la siguiente manera el pensamiento estropeado anteriormente:

$$\begin{array}{ll}
 pr(Lx, Dx) & pr(\sim Lx, Cx) \\
 Da & Ca \\
 \therefore pr(La, pr(Lx, Dx) \text{ y } Da) & \therefore pr(\sim La, pr(\sim Lx, Cx) \text{ y } Ca)
 \end{array}$$

Si queremos predecir el tiempo, estaremos especialmente interesados en:

$$pr(\sim La, e) \text{ o } pr(La, e)$$

donde  $e$  es toda la evidencia pertinente que tenemos. Pero es claro que no podemos inferir ninguno de éstos de los dos argumentos anteriores, incluso si  $e$  es simplemente la conjunción de sus premisas (e incluso si sustituimos nuestra « $pr$ » cualitativa por una medida numérica de grado de apoyo).

Me propongo aplicar el modelo al razonamiento práctico de la manera obvia. La idea central es que un principio moral como «mentir es (*prima facie*) erróneo» no puede tratarse coherentemente como un condicional cuantificado universalmente, sino que ha de reconocerse que significa algo como «que un acto sea una mentira lo hace, *prima facie*, erróneo»; en símbolos, « $pf(Ex, Mx)$ ». El concepto de *prima facie*, tal como se necesita en filosofía moral, relaciona proposiciones. En la gramática lógica, «*prima facie*» no es un operador de oraciones singulares y mucho menos de predicados de acción, sino de pares de oraciones relacionadas como (si expresaran) juicios y fundamentos morales. He aquí cómo se vería el razonamiento práctico, tergiversado por ( $M_1$ ), ( $m_1$ ) y ( $C_1$ ), cuando se lo reconstituye:

( $M_6$ )  $pf(x$  es mejor que  $y$ ,  $x$  es una abstención de fornicación y  $y$  es un acto de fornicación)

( $m_6$ )  $a$  es una abstención de fornicación y  $b$  es un acto de fornicación

$\therefore (C_6) pf(a$  es mejor que  $b$ , ( $M_6$ ) y ( $m_6$ ))

De manera similar, ( $M_2$ ) y ( $m_2$ ), cuando se reescriben del modo nuevo y se rotulan ( $M_7$ ) y ( $m_7$ ), darán:

(C<sub>7</sub>)  $pf(b \text{ es mejor que } a, (M_7) \text{ y } (m_7))$

Un juicio en el que tendremos especial interés es:

(C<sub>8</sub>)  $pf(a \text{ es mejor que } b, e)$

donde  $e$  son todas las consideraciones pertinentes que nos son conocidas incluyendo, al menos, ( $M_6$ ), ( $m_6$ ), ( $M_7$ ) y ( $m_7$ ).

Claro está que (C<sub>8</sub>) no se sigue lógicamente de nada de lo que pasó antes, pero a este respecto el razonamiento moral no parece ser peor que predecir el tiempo. En ninguno de estos casos conocemos una fórmula general para computar si una conjunción apoya, o hasta qué punto lo hace, una conclusión, a partir de si la apoya cada conyunto o a partir de hasta qué punto lo hacen. A este respecto, tampoco hay pérdida alguna en nuestra estrategia de relativizar los juicios morales; no tenemos pista alguna acerca de cómo llegar a (C<sub>8</sub>) a partir de razones, pero su prototipo defectuoso, (C<sub>3</sub>), no se encontraba en mejor forma. Sin embargo, ha habido una pérdida de pertinencia, pues la condicionalización que impide que (C<sub>6</sub>) choque con (C<sub>7</sub>) y que (C<sub>8</sub>) choque con cualquiera de éstos también aísla de la acción a los tres. Al defender P1 y P2 he argumentado que la acción intencional está directamente dirigida a juicios incondicionales como «sería mejor hacer  $a$  que hacer  $b$ ». El razonamiento que se detiene en juicios condicionales tales como (C<sub>8</sub>) es práctico sólo en su tema, pero no en su resultado.<sup>22</sup>

Sin embargo, el razonamiento práctico con frecuencia llega a juicios incondicionales de que una acción es mejor que otra —de otra manera, no habría una cosa tal como actuar conforme a una razón. Los elementos mínimos de tal razonamiento son éstos: el agente acepta alguna razón (o conjunto de razones)  $r$  y sostiene que  $pf(a \text{ es mejor que } b, r)$  y ésta constituye la razón por la que juzga que  $a$  es mejor que  $b$ . Bajo estas condiciones, el agente hará  $a$  si es que hace  $a$  o  $b$  intencionalmente y su razón para hacer  $a$  más bien que  $b$  será idéntica a la razón por la que juzga que  $a$  es mejor que  $b$ .

Esta versión modificada de actuar conforme a una razón no toca ni a P1 ni a P2 y sigue siendo convincente la observación de Aris-

22. En el ensayo 5 se continúa el examen de esta aseveración.



tóteles de que la conclusión (de un razonamiento práctico) es una acción. Pero ahora no hay ninguna dificultad (lógica) en el hecho de la incontinencia, pues al *ákrates* se lo caracteriza como alguien que sostiene que, hechas todas las consideraciones, sería mejor hacer *b* que hacer *a*, aun cuando él hace *a* más bien que *b* y con una razón. La dificultad lógica se ha desvanecido porque un juicio de que *a* es mejor que *b*, hechas todas las consideraciones, es un juicio relacional o *pf* y, así, no puede entrar en conflicto lógico con un juicio incondicional.

Posiblemente se aceptará que P1–P3, tal como aquí se interpretan, no producen una contradicción. Pero, al mismo tiempo, puede surgir una duda sobre si P3 es plausible, dada esta interpretación, pues ¿cómo es posible que un hombre juzgue que *a* es mejor que *b*, hechas todas las consideraciones, y no juzgue que *a* es mejor que *b*?

Rápidamente se descarta una confusión potencial. «*a* es mejor que *b*, hechas todas las consideraciones (esto es, acerca de todas las verdades, morales y no)» ciertamente implica formalmente «*a* es mejor que *b*», y no queremos explicar la incontinencia como una mera falla lógica. La frase «hechas todas las consideraciones» debe referirse, claro está, sólo a aquello que el agente cree o sostiene, la suma de sus principios, opiniones, actitudes y deseos pertinentes. Empero, dejar esto claro parece que sólo subraya la dificultad real. Ahora deseamos preguntar: ¿cómo es posible que un hombre juzgue que *a* es mejor que *b* sobre la base de que *r* y, sin embargo, no juzgue que *a* es mejor que *b* cuando *r* es la suma de todo lo que le parece pertinente? Cuando decimos que *r* contiene todo lo que al agente le parece pertinente, ¿no queremos decir, tan sólo, que no se ha omitido nada que pueda influir en su juicio de que *a* es mejor que *b*?

Puesto que lo que es central para la solución propuesta en este escrito sobre el problema de la incontinencia es el contraste entre juicios (*prima facie*) evaluativos condicionales y juicios evaluativos sin más, quizás podamos dar una caracterización de la incontinencia que evite la problemática expresión «hechas todas las consideraciones». Una modificación plausible de nuestra definición original de incontinencia, (D), podría etiquetar una acción, *x*, como incontinente con sólo que, simplemente, el agente tenga una mejor razón

para hacer cualquier otra cosa; él hace  $x$  por una razón  $r$ , pero tiene una razón  $r'$  que incluye  $r$  y más, con base en la cual juzga que  $y$  es una alternativa mejor que  $x$ .<sup>23</sup> Claro está que podría también haber sido incontinente de su parte haber hecho  $y$ , puesto que podría haber tenido una razón aún mejor para realizar alguna tercera acción  $z$ . Siguiendo esta línea de pensamiento, podríamos decir que una acción  $x$  es continente si  $x$  se hace por una razón  $r$  y no hay ninguna razón  $r'$  (que incluye a  $r$ ), con base en la cual el agente juzga alguna acción mejor que  $x$ .

Esto muestra que podemos darle sentido a la incontinencia sin apelar a la idea de la sabiduría total de un agente  $y$ , de cualquier manera, la nueva formulación podría considerarse una mejora sobre (D), ya que permite (creo que de manera correcta) que haya acciones incontinentes incluso cuando no se hace ningún juicio a la luz de todas las razones. Sin embargo, no podemos excluir el caso en el que se hace un juicio a la luz de todas las razones, así que puede pensarse que aún permanece la dificultad subyacente.

De hecho, sin embargo, la dificultad no es real. Todo juicio se hace a la luz de todas las razones en este sentido: que se hace en presencia de esa totalidad y está condicionado por ella. Pero esto no significa que todo juicio sea razonable o que el agente piense que lo sea, con base en esas razones, ni que se llegó al juicio a partir de esa base mediante un proceso de razonamiento. No hay ninguna paradoja en suponer que una persona sostiene, en ocasiones, que todo lo que cree y aprecia apoya cierto curso de acción cuando, al mismo tiempo, esas mismas creencias y valores son lo que causa que él rechace esa vía de acción.<sup>24</sup> Si  $r$  es la razón que alguien tiene para sostener que  $p$ , entonces pienso que el que sostenga  $r$  debe ser una causa de que sostenga que  $p$ . Pero, y esto es lo que es crucial aquí, el que sostenga que  $r$  puede ser la causa de que sostenga que  $p$ , sin que  $r$  sea su razón; de hecho, el agente puede incluso pensar que  $r$  es una razón para rechazar  $p$ .

23. Amparados por esta formulación, podríamos querer excluir el caso en el que el agente hace aquello para lo que tiene la mejor razón, pero no lo hace por esa razón.

24. En este lugar, me parece que mi versión de la incontinencia es muy cercana a la de Aristóteles. Véase G.E.M. Anscombe, «Thought and Action in Aristotle».

Entonces, es posible ser incontinente incluso si son verdaderos P1 y P2. Pero, conforme a este análisis, ¿cuál es la falta en la incontinencia? Como es claro ahora, el *ákrates* no sostiene creencias lógicamente contradictorias ni su falla es, necesariamente, una falla moral. Lo que está mal es que el hombre incontinente actúa y juzga irracionalmente, pues esto, seguramente, es lo que debemos decir de un hombre que va en contra de su propio mejor juicio. Carnap y Hempel han argumentado que hay un principio que no es ninguna parte de la lógica del razonamiento inductivo (o estadístico), pero que es una directiva que aceptará el hombre racional. Éste es el requisito de *evidencia total a favor del razonamiento inductivo*: da crédito a la hipótesis apoyado en toda la evidencia pertinente disponible.<sup>25</sup> Sugiero que hay un principio análogo que aceptará el hombre racional al aplicar el razonamiento práctico: realiza la acción que se juzga que es la mejor con base en todas las razones pertinentes disponibles. Sería apropiado denominar esto el *principio de continencia*. Puede parecer algo extraño hacer del requisito de evidencia total un imperativo (¿puede uno formar las propias creencias sobre pedido?), pero no hay ninguna extrañeza así con respecto al principio de continencia. Éste nos exhorta a acciones que podemos realizar si lo queremos; nos deja los motivos a nosotros. Lo que es difícil es adquirir la virtud de la continencia, hacer del principio de continencia algo propio. Pero, en principio, no hay ninguna razón por la que sea algo más difícil hacerse continente que hacerse casto o valiente. Se obtiene un sentido vívido de las dificultades en la extraordinaria plegaria de San Agustín: «Dame la castidad y la continencia, sólo que todavía no.» (*Confesiones*, VIII, vii).

¿Por qué jamás realizaría alguien cualquier acción cuando pensase que, hechas todas las consideraciones, sería mejor otra acción? Si esto es una solicitud de explicación psicológica, entonces, sin duda, las respuestas se referirán a los interesantes fenómenos que nos son familiares en la mayoría de los estudios sobre la incontinencia: autoengaño, deseos avasallantes, falta de imaginación y todo lo demás. Pero si la pregunta se lee como cuál es la razón del agente para hacer *a* cuando cree que sería mejor, hechas todas las considera-

25. Véase Hempel, *op. cit.*, pp. 397-403, para modificaciones importantes y otras referencias.

ciones, hacer otra cosa, entonces la respuesta debe ser: para esto, el agente no tiene ninguna razón.<sup>26</sup> Nosotros percibimos una criatura como racional en la medida en que somos capaces de ver sus movimientos como parte de un modelo racional que también comprende pensamientos, deseos, emociones y voliciones (en esto, mucho nos ayudan las acciones que concebimos como emisiones lingüísticas). Por inferencias defectuosas, evidencia incompleta, falta de diligencia o interés débil, con mucha frecuencia no logramos detectar un patrón que se encuentra ahí. Pero, en el caso de la incontinencia, el intento por encontrar la razón en la conducta está necesariamente sujeto a cierta frustración.

Lo que la incontinencia tiene de especial es que el actor no puede entenderse a sí mismo: él reconoce, en su propia conducta intencional, algo esencialmente irracional.

26. Claro está que tiene una razón para hacer *a*; lo que le falta es una razón para no dejar que prevalezca su mejor razón para no hacer *a*.

### 3. DE LA ACCIÓN

¿Qué sucesos en la vida de una persona revelan actuación; cuáles son sus actos y sus obras, a diferencia de meros acontecimientos en su historia; cuál es la marca que distingue sus acciones?

Esta mañana me despertó el sonido de alguien practicando el violín. Dormité un rato, luego me levanté, me lavé, me rasuré, me vestí y bajé las escaleras, y apagué, de paso, la luz del corredor. Me serví un poco de café, me tropecé con el borde de la alfombra del comedor y derramé el café cuando buscaba el *New York Times*.

Algunos de estos detalles registran cosas que yo hice; otros, cosas que me acontecieron, cosas que me sucedieron cuando iba hacia el comedor. Entre las cosas que hice están el levantarme, lavarme, rasurarme, bajar las escaleras y derramar mi café. Entre las cosas que me sucedieron están el ser despertado y el tropezar con el borde de la alfombra. Un caso límite fue, quizás, el dormir. Podrían introducirse dudas acerca de otros casos si se detalla más la historia. El tropezón puede ser deliberado y, cuando lo sea, se considerará como una cosa hecha. Yo podría haber apagado la luz al pegarme inadvertidamente al interruptor; ¿habría sido, entonces, un acto mío, o siquiera algo que yo hice?

Resolver muchos ejemplos está al alcance de la mano y esto apoya la esperanza de que hay un principio interesante en operación, un principio que, de hacerse explícito, podría ayudar a explicar por qué son difíciles los casos difíciles. Por otra parte, muchos casos hacen surgir dificultades. La cuestión misma parece quedar fuera de foco cuando comenzamos a escudriñar frases tales como «lo que hizo», «sus acciones», «lo que le sucedió» y, con frecuencia, es

pertinente para la propiedad de la respuesta la forma que le demos a la pregunta (despertar fue, quizás, algo que yo hice, pero no una acción). Hemos de mantener un sentido vívido de la posibilidad de que la pregunta con la que comenzamos, como lo sugirió Austin, esté desencaminada.<sup>1</sup>

Sin embargo, en este ensayo pongo una vez más a prueba el supuesto positivo de que la pregunta es una buena pregunta, de que hay una subclase de sucesos bastante bien definida que son acciones. Los costos de ese supuesto son los usuales: simplificación excesiva, dejar de lado una amplia clase de excepciones, pasar por alto las distinciones a las que apuntan la gramática y el sentido común, apelar a una legislación lingüística disfrazada. Con suerte, aprendemos algo de tales métodos. Después de todo, puede haber importantes verdades generales en esta área y, si las hay, ¿de qué otra manera las descubriremos?

Con frecuencia los filósofos parecen pensar que debe haber algún indicador gramatical simple de actuación, pero no se ha descubierto ninguno. Yo drogué al centinela, yo contraí la malaria, yo bailé, yo me desmayé, yo pateé a Ramírez, yo viví más que Pérez; ésta es una serie de ejemplos que intentan mostrar que una persona nombrada como sujeto en oraciones activas (sea el verbo transitivo o no), o como objeto en oraciones pasivas, puede o puede no ser el agente del suceso registrado.<sup>2</sup>

Otro error común es pensar que puede hacerse una lista de verbos conforme a si éstos imputan o no actuación a un sujeto o a un objeto. Lo que provoca el error es que esto es verdadero de algunos verbos. Decir que una persona ha cometido un error, ha insultado a su tío o ha hundido el *Bismarck*, es automáticamente acusarla de ser la autora de esos sucesos; y mencionar a alguien en la posición de sujeto de una oración con el verbo en voz pasiva es, hasta donde puedo ver, asegurar que no es el agente de la acción. Pero, con mucha frecuencia, una oración registrará un episodio en la vida de un agente y nos dejará a oscuras acerca de si fue una acción. He aquí algunos ejemplos: él parpadeó, rodó fuera de la cama, prendió la luz,

1. Véase el artículo de J.L. Austin, «A Plea for Excuses», pp. 126–127.

2. La propuesta se desarrolla en el artículo de Irving Thalberg, «Verbs, Deeds and What Happens to Us», pp. 259–260.

tosió, bizqueó, sudó, derramó el café y se tropezó con la alfombra. Sabremos si estos sucesos son acciones sólo hasta que sepamos más que lo que nos dice el verbo. Al considerar la información adicional que podría resolver el asunto, podríamos encontrar una respuesta a la pregunta acerca de qué es lo que hace que una parte de la biografía sea una acción.

Di una pista en mi diario fragmentario del principio. Tropezarse con la alfombra no es normalmente una acción, pero lo es si se hace intencionalmente. Quizás, entonces, ser intencional es el rasgo distintivo pertinente. Si lo fuese, ayudaría a explicar por qué algunos verbos implican actuación, pues algunos verbos describen acciones que no pueden ser sino intencionales. Ejemplos de esto son aseverar, hacer trampa, sacar una raíz cuadrada y mentir.

Sin embargo, este rasgo no es suficiente, pues aun cuando la intención implique actuación, la conversa de esto no vale. Así, derramar el café, hundir el *Bismarck* e insultar a alguien son todas cosas que pueden o no hacerse intencionalmente; pero incluso cuando no son intencionales, son normalmente acciones. Si, por ejemplo, derramo intencionalmente el contenido de mi taza, pensando erróneamente que es té cuando es café, entonces derramar el café es algo que yo hago, es una acción mía, aun cuando no la haga intencionalmente. Por otra parte, si derramo el café porque me mueves la mano, no se me puede denominar agente. Sin embargo, aun cuando pueda apresurarme a pedir excusas, ni siquiera en este caso es incorrecto decir que yo derramé el café. Así, debemos distinguir tres situaciones en las que es correcto decir que derramé el café: en la primera, lo hago intencionalmente; en la segunda, no lo hago intencionalmente, pero es mi acción (pensé que era té); en la tercera, de ninguna manera es mi acción (moviste mi mano).<sup>3</sup>

Ciertos tipos de error son particularmente interesantes: leer mal un signo, interpretar mal una orden, subestimar un peso o calcular mal una suma. Hablando estrictamente, estas cosas no pueden hacerse intencionalmente. Uno puede fingir que lee mal un signo, se puede subestimar un peso por pereza o por falta de atención,

3. Esta división tripartita no ha de confundirse con el sutil trabajo de Austin sobre la diferencia entre propósito, intención y deliberación, en «Three Ways of Spilling Ink».

o escribir deliberadamente la que uno sabe que es una respuesta incorrecta a una suma; pero ninguna de éstas es una equivocación intencional. Cometer un error de uno de los tipos mencionados es fallar en hacer lo que uno pretende y, dejando de lado la paradoja freudiana, uno no puede tener la intención de fallar. Entonces, estos errores no son intencionales y, sin embargo, son acciones. Para ver esto sólo debemos darnos cuenta de que cometer un error debe ser, en cada caso, hacer alguna otra cosa intencionalmente. Una mala lectura debe ser una lectura, aun cuando uno no logre lo que quería; malinterpretar una orden es un caso de interpretarla (y con la intención de hacerlo correctamente); subestimar es estimar y un mal cálculo es un cálculo (aun cuando sea uno que fracase).

¿Podemos señalar ahora qué sucesos comprenden la actuación? Lo hacen las acciones intencionales y, también, algunas otras cosas que hacemos. ¿Cuál es el elemento común? Consideremos, nuevamente, derramar el café. Yo soy el agente si derramo el café con el propósito de derramar el té, pero no si me mueves la mano. ¿Cuál es la diferencia? Parece que la diferencia se encuentra en el hecho de que en un caso, pero no en el otro, yo, intencionalmente, estoy haciendo *algo*. Mi derramar el contenido de mi taza fue intencional; sucede que este mismo acto puede redesccribirse como mi derramar el café. Claro está que, redescrita de esta manera, la acción ya no es intencional; pero, aparentemente, este hecho no es pertinente para la cuestión de la actuación.

Y así, según pienso, tenemos una respuesta correcta para nuestro problema: un hombre es el agente de un acto si lo que hace puede redesccribirse bajo un aspecto que lo haga intencional.

Lo que hace posible esta respuesta es la opacidad semántica, o intensionalidad, de las atribuciones de intención. Hamlet mata intencionalmente al hombre detrás del tapiz, pero no mata intencionalmente a Polonio. Sin embargo, Polonio es el hombre detrás del tapiz y, así, el matar de Hamlet al hombre detrás del tapiz es idéntico a su matar a Polonio. Es un error suponer que hay una clase de acciones intencionales; si tomásemos esta dirección, estaríamos orillados a decir que una y la misma acción fue, a la vez, intencional y no intencional. Como un primer paso para enderezar las cosas podemos intentar hablar, no de acciones, sino, en su lugar, de oraciones y de descripciones de acciones. En el caso de la actuación, mi



propuesta podría entonces formularse así: una persona es el agente de un suceso si y sólo si hay una descripción de lo que ella hizo que haga verdadera una oración que dice que lo hizo intencionalmente. Esta formulación, con su cuantificación sobre entidades lingüísticas, no puede considerarse completamente satisfactoria. Pero, para mejorarla, se requeriría un análisis semántico acerca de actitudes proposicionales.<sup>4</sup>

Dejando de lado la necesidad de otros refinamientos, el criterio de acciones propuesto parece adecuarse a los ejemplos que hemos estudiado. Supóngase que un oficial apunta un torpedo a un barco que piensa que es el *Tirpitz* y hunde realmente el *Bismarck*. Entonces hundir el *Bismarck* es su acción, pues esa acción es idéntica a su intento de hundir el barco que consideraba que era el *Tirpitz*, lo cual es intencional. De manera similar, derramar el café es el acto de una persona que lo hace al derramar intencionalmente el contenido de su taza. También ahora es más claro por qué los errores son acciones, pues cometer un error debe ser hacer algo con la intención de alcanzar un resultado que no se presenta.

Si, como insisto, podemos decir que una persona hace como agente cualquier cosa que haga intencionalmente bajo una descripción, entonces aun cuando el *criterio* de actuación es intensional en el sentido semántico, es puramente extensional la *expresión* misma de la actuación. La relación que se da entre una persona y un suceso, cuando el suceso es una acción realizada por la persona, vale sin importar cómo se describan los términos. Por tanto, podemos hablar, sin confusión, de la clase de los sucesos que son acciones, lo cual no podemos hacer de las acciones intencionales.

Podría pensarse que el concepto de una acción es, irremediablemente, algo confuso debido a lo difícil que es decidir si, por ejemplo, es o no una acción golpear a un policía, caerse por las escaleras o desinflar el ego de una persona. Pero si ser una acción es un rasgo que tienen los sucesos particulares, independientemente de cómo se los describa, no hay ninguna razón que nos haga esperar, en general, que seamos capaces de determinar, tan sólo por conocer algún rasgo

4. Para un intento de tal teoría, véase mi «On Saying That» y la discusión en el ensayo 6.

de ese suceso (que, por ejemplo, es un caso de golpear a un policía) si es o no es una acción.

¿Es tan amplio nuestro criterio que incluirá como acciones sucesos que normalmente no se considerarían como tales? Por ejemplo, tropezarse con el borde de la alfombra ¿no es tan sólo parte de mi caminar intencional hacia el comedor? Pienso que no. Un movimiento intencional mío fue la causa de que me tropezara y, así, yo me tropecé; ésta fue una acción, aun cuando no fuera intencional. Pero «yo tropecé» y «yo hice que me tropezara» no dan cuenta del mismo suceso. La segunda oración implica formalmente la primera, porque tropezarme es hacer algo que tiene como resultado el que yo tropiece; pero, claro está, hacer algo que tiene como resultado el que yo tropiece no es idéntico a lo que causa.

La extensionalidad de la expresión de actuación sugiere que el concepto de actuación es más simple o más básico que el de intención pero, desafortunadamente, el camino que hemos recorrido no muestra cómo explotar la pista, pues todo lo que hasta ahora tenemos es una forma de distinguir casos de actuación apelando a la noción de intención. Esto es analizar lo oscuro apelando a lo más oscuro —no es un proceso tan inútil como a menudo se piensa que es, pero sigue siendo decepcionante. Debemos intentar ver si podemos encontrar una marca de la actuación que no use el concepto de intención.

La noción de causa puede proporcionar la clave. Con respecto a la causación hay cierta simetría aproximada entre intención y actuación. Si digo que Pérez incendió la casa para cobrar el seguro, explico su acción, en parte, dando una de sus causas, a saber, su deseo de cobrar el seguro. Si digo que Pérez convirtió en cenizas la casa por prender fuego a la colcha, entonces explico el incendio dando una causa, a saber, la acción de Pérez. En ambos casos, la explicación causal toma la forma de una descripción más completa de una acción, sea en términos de una causa o de un efecto. Describir una acción como algo que tiene cierto propósito o un resultado buscado es describirla como un efecto; describirla como una acción que tiene cierto resultado es describirla como una causa. Típicamente, las atribuciones de intención son excusas y justificaciones; las atribuciones de acciones son típicamente acusaciones o asignaciones de responsabilidad. Claro está que los dos tipos de atribuciones no se excluyen mutuamente, puesto que dar la intención con la que se hizo

un acto también es, y necesariamente, atribuir actuación. Si Bruto asesinó a César con la intención de eliminar a un tirano, entonces una causa de su acción fue un deseo de eliminar a un tirano y un efecto fue la muerte de César. Si el oficial hundió el *Bismarck* con la intención de hundir el *Tirpitz*, entonces una acción suya la causó su deseo de hundir el *Tirpitz* y tuvo como consecuencia que se hundió el *Bismarck*.<sup>5</sup>

Estos ejemplos y otros más sugieren que, en cada caso de acción, el agente hace que suceda, provoca, produce o realiza el suceso del que fue agente, y estas frases, a su vez, parecen fundadas en la idea de causa. Entonces ¿podemos decir que ser el autor o agente de un suceso es causarlo? Esta tesis, o alguna aparentemente muy similar a ésta, la han propuesto o asumido diversos autores recientes.<sup>6</sup> Así que hemos de considerar si introducir de esta manera la noción de causación puede mejorar nuestra comprensión del concepto de actuación.

Es claro que puede hacerlo, al menos hasta cierto punto, pues una manera importante de justificar una atribución de actuación es mostrando que algún suceso fue causado por algo que el agente hizo. Si yo envenené la toronja matutina de alguien con la intención de matarlo, y tuve éxito, entonces yo causé su muerte al poner veneno en su comida y es por eso que soy el agente de su asesinato. Cuando me las arreglo para ofender a alguien al denigrar su corbata, yo causo la ofensa, pero es otro suceso, las palabras malévolas que dije, el que es la causa de la ofensa.

La noción de causa a la que aquí se ha apelado es la causalidad ordinaria de sucesos, la relación, cualquiera que sea, que se da entre dos sucesos cuando uno es la causa del otro. Pues aun cuando digamos que el agente causó la muerte de la víctima, esto es, que él la

5. En el ensayo 1 desarrollo el tema de que dar una razón o una intención con la que se realiza una acción es, entre otras cosas, describir la acción en términos de una causa. En este ensayo exploro cómo los efectos de las acciones entran en nuestras descripciones de ellas.

6. Por ejemplo Roderick Chisholm, «Freedom and Action»; Daniel Bennett, «Action, Reason and Purpose»; Anthony Kenny, *Action, Emotion and Will*; Georg Henrik von Wright, *Norm and Action*; Richard Taylor, *Action and Purpose*. Otras críticas de este tipo de análisis causal de la actuación pueden encontrarse en Irving Thalberg, «Do We Cause our Own Actions?», y en el ensayo 6.

mató, ésta es una forma elíptica de decir que algún acto del agente —algo que él hizo, tal como poner veneno en la toronja— causó la muerte de la víctima.

No todo suceso que atribuimos a un agente puede explicarse como algo causado por otro suceso del cual él es el agente; algunos actos deben ser primitivos en el sentido de que no pueden analizarse en términos de sus relaciones causales con actos del mismo agente. Pero entonces la causalidad de sucesos no puede usarse de esta manera para explicar la relación entre un agente y una acción primitiva. La causalidad de sucesos puede extender la responsabilidad de una acción a las consecuencias de la acción, pero no puede ayudar a explicar la primera atribución de actuación de la que dependen las demás.<sup>7</sup>

Si interpretamos generosamente la idea de un movimiento corpóreo, podemos argumentar a favor de decir que todas las acciones primitivas son movimientos corpóreos. La generosidad debe ser lo suficientemente abierta como para abarcar «movimientos» tales como permanecer inmutable y actos mentales como decidir y computar. No me propongo examinar ahora estos ejemplos difíciles; si estoy equivocado acerca del alcance preciso de las acciones primitivas, esto no afectará mi argumento principal. Sin embargo, es importante mostrar que en acciones ordinarias tales como señalar con un dedo o amarrarse las cintas de los zapatos, la acción primitiva es un movimiento corpóreo.

Puedo imaginar al menos dos objeciones a esta aseveración. Primero, puede decirse que a fin de señalar con mi dedo, hago algo que causa que se mueva el dedo, a saber, contraer algunos músculos, y quizás esto requiera que yo haga que se lleven a cabo ciertos sucesos en mi cerebro. Pero estos sucesos no parecen ser movimientos corpóreos ordinarios. Pienso que las premisas de este argumento pueden ser verdaderas, pero que la conclusión no se sigue. Puede ser

7. Aquí, y en lo que sigue, supongo que hemos dejado de lado un análisis de la actuación que comience analizando el concepto de intención, de actuar con una intención o de una razón al actuar. Estos conceptos pueden analizarse, al menos en parte, en términos de causalidad de sucesos. En el ensayo I intento mostrar que aun cuando las creencias y los deseos (y estados mentales similares) no son sucesos, podemos decir con propiedad que son causas de acciones intencionales y, cuando decimos esto, nos apoyamos en el concepto ordinario de causalidad de sucesos.

verdad que causo que mi dedo se mueva al contraer ciertos músculos y posiblemente causo que los músculos se contraigan al hacer que acontezca un suceso en mi cerebro. Pero esto no muestra que señalar con mi dedo no sea una acción primitiva, pues no muestra que yo deba hacer alguna otra cosa que la cause. Hacer algo que cause que mi dedo se mueva no me causa que yo mueva mi dedo; *es mover mi dedo.*

Al estudiar ejemplos como éste, Chisholm ha sugerido que, aun cuando pueda decirse que un agente hace que se den ciertos sucesos cerebrales cuando son estos sucesos los que causan que su dedo se mueva, hacer que se den los sucesos cerebrales no puede decirse que sea algo que él hace. Chisholm también piensa que muchas cosas de las que el agente es causa de que sucedan, en el sentido de que son sucesos causados por cosas que él hace, no son sucesos de los que él es el agente. Así, si mover el dedo es algo que hace un hombre, y este movimiento causa que se muevan algunas moléculas de aire, entonces aunque pueda decirse que el hombre causó que se movieran las moléculas y, por tanto, que movió las moléculas, esto no es algo que él hizo.<sup>8</sup>

No me parece que ésta sea una distinción clara o útil; todos los casos que presenta Chisholm de hacer que algo suceda, hasta donde llega mi intuición, son casos de actuación, situaciones en las que podemos conceder, y concedemos, que la persona hizo algo, que fue un agente. Cuando una persona hace que acontezca un suceso en su cerebro, normalmente no sabe que está haciendo esto, y Chisholm parece sugerir que, por esta razón, no podemos decir que es algo que ella hace. Pero incluso un hombre puede estar haciendo algo intencionalmente y no saber que lo está haciendo; así que, claro está, puede estar haciéndolo sin saber que lo está haciendo (un hombre puede estar haciendo diez copias al carbón mientras escribe y esto puede ser intencional; sin embargo, él puede no saber que lo está haciendo; todo lo que sabe es que lo está intentando).

La acción requiere que lo que el agente haga sea intencional bajo alguna descripción, y pienso que esto requiere, a su vez, que lo que el agente haga sea algo conocido por éste bajo alguna descripción. Pero nuestros ejemplos cumplen con esta condición. Un hombre que

8. Chisholm, «Freedom and Action».

alza el brazo tanto pretende hacer con su cuerpo lo que se necesite para hacer que su brazo suba, como sabe que lo está haciendo. Y, claro está, los sucesos cerebrales y los movimientos de los músculos son precisamente lo que se necesita. Así pues, aun cuando el agente pueda no saber los nombres o las localizaciones de los músculos pertinentes, y ni siquiera saber que tiene un cerebro, lo que hace que suceda en su cerebro y en sus músculos cuando mueve el brazo es, bajo una descripción natural, algo que él pretende hacer y de lo que tiene conocimiento.

La segunda objeción a la aseveración de que las acciones primitivas son movimientos corpóreos proviene de la dirección opuesta: afirma que algunas acciones primitivas abarcan más que un movimiento del cuerpo. Cuando ato las cintas de mis zapatos, por una parte hay un movimiento de mis dedos y, por otra, un movimiento de las cintas. Pero ¿es posible separar estos sucesos llamando sólo al primero mi acción? Lo que hace que la separación sea un problema es que no me parece que yo sea capaz de describir o de pensar cómo muevo mis dedos, aparte de mover las cintas. Yo no muevo mis dedos cuando intento causar que mis zapatos queden atados, ni soy capaz de mover mis dedos de la manera apropiada cuando no hay cintas presentes (éste es un truco que yo podría aprender). De manera similar, se podría argumentar que cuando la mayoría de la gente emite palabras, lo hace sin saber qué músculos mover o cómo poner su boca a fin de emitir las palabras que quiere; así que aquí, nuevamente, parece que una acción primitiva debe incluir más que un movimiento corpóreo, a saber, un movimiento del aire.

La objeción fracasa por la misma razón que la anterior. Todo depende de si hay o no una descripción adecuada de la acción. Se asume correctamente que a menos que el agente mismo sea consciente de lo que él está haciendo sólo con su cuerpo, a menos que pueda concebir sus movimientos como un suceso separado físicamente de cualquier otra cosa que suceda, sus movimientos corpóreos no pueden ser su acción. Pero se supone, erróneamente, que tal darse cuenta y tal concepción son imposibles en el caso de hablar o en el de amarrarse las propias cintas de los zapatos. Pues un agente siempre sabe cómo mueve su cuerpo cuando, al actuar intencionalmente, mueve su cuerpo, en el sentido de que hay *alguna* descripción del movimiento bajo la que él sabe que lo hace. Ciertamente, tales descripciones pueden

ser triviales y poco reveladoras; esto es lo que asegura su existencia. Así, si yo ato mis cintas de los zapatos, he aquí una descripción de mis movimientos: muevo mi cuerpo justamente de la manera que se requiere para que yo ate las cintas de mis zapatos. De manera similar, cuando emito palabras, es verdad que soy incapaz de describir lo que hacen mi lengua y mi boca o de nombrar los músculos que muevo. Pero no necesito la terminología del terapeuta del lenguaje; lo que hago es mover mi boca y mis músculos, como sé hacerlo, justamente de la manera que se necesita para producir las palabras que tengo en mente.

Así que, después de todo, no hay ningún problema para producir descripciones familiares y correctas de mis movimientos corpóreos, y éstos son los sucesos que causan otros sucesos tales como el que las cintas de mis zapatos estén siendo atadas o que el aire vibre con mis palabras. Claro está que la astucia descriptiva consiste en describir las acciones como los movimientos con los efectos correctos; pero esto no muestra que la astucia no haya servido. Lo que se necesitaba no era una descripción que no mencionase los efectos, sino una descripción que se adecuase a la causa. Concluyo que no hay obstáculo alguno para decir que nuestras acciones primitivas, al menos si hacemos de lado casos problemáticos tales como los actos mentales, son movimientos corpóreos.

Volvamos a la cuestión acerca de si puede analizarse el concepto de acción en términos del concepto de causalidad; lo que nuestro estudio ha mostrado es que podemos concentrarnos en acciones primitivas. La noción ordinaria de causalidad de sucesos es útil para explicar cómo la actuación puede extenderse de acciones primitivas a acciones que se describen de otras maneras, pero no puede, de la misma manera, explicar el sentido básico de actuación. Entonces, lo que debemos preguntar es si hay acaso otro tipo de causalidad que no se reduzca a causalidad de sucesos, y que nos auxilie para entender la actuación. Podemos (siguiendo a Thalberg) denominar a este tipo de causalidad *causalidad de agentes*.

Si, por las razones que acabamos de dar, nos restringimos a las acciones primitivas, ¿qué tan bien da cuenta la idea de causalidad de agentes de la relación entre un agente y su acción? Tenemos este dilema: o bien el que un agente cause una acción primitiva es un suceso separado de la acción primitiva, en cuyo caso tenemos los

problemas acerca de los actos de la voluntad o algo peor, o no está separado, en cuyo caso no parece que haya ninguna diferencia en decir que alguien causó una acción primitiva y en decir que él fue el agente.

Tomemos el primer término: supongamos que causar una acción primitiva (en el sentido de causalidad de agentes) introduce un suceso separado de la acción, y presumiblemente anterior a ella. Este suceso anterior, a su vez, debe o bien ser una acción o bien no serlo. Si es una acción, entonces la acción con la que comenzamos, en contra de nuestro supuesto, no era primitiva. Si no es una acción, entonces hemos intentado explicar la actuación apelando a una noción aún más oscura, la de un causar que no es un hacer.

Uno se atora en el segundo término del dilema si supone que la causación de agentes *no* introduce un suceso además de la acción primitiva, pues entonces ¿qué más hemos dicho cuando decimos que el agente causó la acción que cuando decimos que fue el agente de la acción? Parece que el concepto de *causa* no desempeña ningún papel. Podemos no detectar la vacuidad de esta sugerencia porque la causalidad, como lo hemos notado, se introduce conspicuamente en las versiones de la actuación; pero, cuando lo hace, es la variante común de la causalidad la que no ilumina la relación entre el agente y sus acciones primitivas.

Explicamos una ventana rota diciendo que un ladrillo la rompió; el poder explicativo que tiene la observación se deriva del hecho de que primero podemos ampliar la explicación de la causa de tal manera que abarque un suceso, el movimiento del ladrillo, y luego podemos acumular evidencia a favor de la existencia de una ley que conecte sucesos tales como los movimientos de objetos rígidos de tamaño medio y el rompimiento de ventanas. La noción ordinaria de causa es inseparable de esta forma elemental de explicación. Pero el concepto de causación de agentes carece por completo de estos rasgos. Lo que distingue la causación de agentes de la causación ordinaria es que no es posible ninguna expansión a una narración de dos sucesos y no se vislumbra ninguna ley. Por la misma razón, nada se explica. Por tanto, parece que no hay ninguna buena razón para usar expresiones tales como «causa», «producir», «hacer que suceda», para *iluminar* la relación entre un agente y su acto. No quiero decir que haya algo erróneo con tales expresiones —hay ocasiones



en las que se presentan naturalmente al hablar de actuación. Pero no pienso que por introducirlas progrese hacia la comprensión de la actuación y de la acción.

La causalidad le es central al concepto de actuación, pero es la causalidad ordinaria entre sucesos la que es pertinente, y concierne a los efectos y no a las causas de las acciones (descontando, como antes, la posibilidad de analizar la intención en términos de causalidad). Una forma de mostrar esto es describiendo lo que Joel Feinberg denomina el «efecto de acordeón»,<sup>9</sup> que es un rasgo importante del lenguaje que usamos para describir acciones. Digamos que un hombre mueve su dedo intencionalmente, presionando así el interruptor, y causando que se prenda la luz, que el cuarto se ilumine y que un merodeador se alerte. Este enunciado tiene las siguientes implicaciones formales: el hombre presionó el interruptor, prendió la luz, iluminó el cuarto y alertó al merodeador. Algunas de estas cosas las hizo intencionalmente, algunas otras no; fuera del movimiento del dedo, la intención no es pertinente para las inferencias, e incluso en este caso sólo se la requiere en el sentido de que el movimiento debe ser intencional bajo alguna descripción. En breve, una vez que él ha hecho una cosa (mover un dedo), cada consecuencia nos ofrece un acto; un agente causa lo que causan sus acciones.<sup>10</sup>

El efecto de acordeón no revelará conforme a qué es intencional un acto. Si alguien mueve su boca de tal manera que produce las palabras «tienes el bate al revés», causando así una ofensa a su

9. Joel Feinberg, «Action and Responsibility».

10. La formulación en esta oración es más precisa que en algunos de mis ejemplos. Supóngase que Ramírez intencionalmente causa que Pérez intencionalmente le dispare a Gómez y lo mate. Ciertamente no concluiremos que Ramírez le disparó a Gómez, y podemos o podemos no decir que Ramírez mató a Gómez. Empero, mi formulación es correcta siempre que podamos pasar de «La acción de Ramírez causó la muerte de Gómez» a «Ramírez causó la muerte de Gómez». Claro está que habrá un conflicto si negamos que (en nuestra narración) pueda decirse tanto de Ramírez como de Pérez que causaron la muerte de Gómez y afirmar, al mismo tiempo, la transitividad de la causalidad. Sin embargo, podríamos preservar la fórmula ante una negación de que, dadas las circunstancias, pueda decirse que Ramírez causó la muerte de Gómez, diciendo que, dadas las circunstancias, también se viene abajo la transitividad de la causalidad. Para mayor discusión del asunto, véanse H.L.A. Hart y A.M. Honoré, *Causation in the Law*; Joel Feinberg, «Causing Voluntary Actions» y J.E. Atwell, «The Accordion-Effect Thesis».

compañero, se aplica el efecto de acordeón, pues podemos decir tanto que dijo esas palabras como que ofendió a su compañero. Sin embargo, es posible que no pretendiese mover su boca de tal manera que produjera esas palabras, ni producirlas, ni ofender a su compañero. Pero el efecto de acordeón no es aplicable si no hay ninguna intención presente. Si el oficial presiona un botón pensando que hará sonar una campana que hará venir a un camarero trayéndole una taza de té, pero, de hecho, dispara un torpedo que hunde el *Bismarck*, entonces el oficial hundió el *Bismarck*; pero si cae contra el botón porque una ola le hizo perder el equilibrio, entonces, aunque las consecuencias sean las mismas, no lo consideraremos a él como agente.

El efecto de acordeón se limita a los agentes. Si Ramírez intencionalmente agita un bate que le pega a una pelota, la cual golpea y rompe una ventana, entonces Ramírez no sólo le pegó a la pelota sino que también rompió la ventana. Pero no decimos que el bate o siquiera su movimiento rompió la ventana, aun cuando, claro está, el movimiento del bate causó el rompimiento. Ciertamente concedemos que los objetos inanimados causan o producen diversas cosas —en nuestro ejemplo, la pelota rompió la ventana. Sin embargo, no es éste el efecto de acordeón de la actuación, sino sólo la elipsis de la causalidad de sucesos. La pelota rompió la ventana: es decir, su movimiento causó el rompimiento.

Por tanto, parece que podemos tomar el efecto de acordeón como una marca de la actuación. Es una manera de inquirir si un suceso es un caso de actuación o de preguntar si podemos atribuir su efecto a una persona. Y, por otra parte, siempre que decimos que una persona ha hecho algo, donde claramente lo que mencionamos no es un movimiento corpóreo, lo hemos hecho el agente no sólo del suceso mencionado, sino de algún movimiento corpóreo que lo produjo. En el caso de los movimientos corpóreos, en ocasiones tenemos una manera breve de mencionar a una persona y un suceso y, sin embargo, dejar abierta la cuestión acerca de si ella fue el agente, como en: Pérez se cayó.

El efecto de acordeón es interesante porque muestra que tratamos las consecuencias de las acciones de manera diferente a como tratamos las consecuencias de otros sucesos. Esto sugiere que, después de todo, hay una prueba lingüística muy simple que en ocasiones

revela que tomamos un suceso como si fuera una acción. Pero, como criterio, difícilmente puede considerárselo satisfactorio; sólo funciona en algunos casos y, claro está, no ofrece ninguna clave acerca de qué es lo que hace que una acción primitiva sea una acción.

En este punto abandono la búsqueda de un análisis del concepto de actuación que no apele a la intención y considero una cuestión relacionada, que se ha presentado al tratar la causalidad de agentes y el efecto de acordeón. La nueva pregunta es qué relación tiene un agente con sus acciones que no son primitivas, con las acciones para cuya descripción rebasamos los meros movimientos corpóreos y nos fijamos en las consecuencias, en lo que el agente ha labrado en el mundo más allá de su piel. Suponiendo que entendemos la actuación en el caso de las acciones primitivas, ¿cómo se relacionan exactamente con las demás? Puede parecer que ya se ha respondido la pregunta que ahora deseo formular, pero de hecho no es así. Lo que es claro es la relación entre una acción primitiva, por ejemplo, mover uno de nuestros dedos de cierta manera, y una consecuencia tal como el que se aten las cintas de nuestros zapatos: ésta es la relación de causalidad de sucesos. Pero esto no da una respuesta clara a la pregunta de cómo el movimiento de las manos se relaciona con la acción de atar las cintas de nuestros zapatos, ni tampoco a la pregunta de cómo se relaciona la acción de atar las cintas de nuestros zapatos con el que se aten las cintas de nuestros zapatos. O, para alterar el ejemplo, si Bruto mató a César al apuñalarlo, ¿cuál es la relación entre estas dos acciones, la relación expresada por «al»? Sin duda es verdad que Bruto mató a César porque el apuñalamiento tuvo como resultado la muerte de César; pero aún tenemos ese tercer suceso cuyas relaciones con los otros son poco claras, a saber, el matar mismo.

Es natural suponer que la acción, cuya mención incluye la mención de un resultado, incluye ella misma, de alguna manera, ese resultado. Así, Feinberg dice que la acción de un hombre puede «comprimirse a un mínimo o bien estirarse» por el efecto de acordeón. «Él giró la llave, abrió la puerta, sobresaltó a Ramírez, mató a Ramírez —son todas cosas que podríamos decir que Martínez hizo con un conjunto idéntico de movimientos corpóreos», nos dice Feinberg. Es justamente en esta relación de «hacer con» o «hacer

al» en la que estamos interesados. Feinberg continúa: «Podemos, si queremos, inflar una acción para incluir un efecto.»<sup>11</sup> Inflar, comprimir, estirar, suenan como operaciones que se realizan sobre uno y el mismo suceso; empero, si, como parece claro, estas operaciones cambian la extensión temporal del suceso, entonces no puede ser uno y el mismo suceso; conforme a la teoría de Feinberg, la acción de abrir la puerta no puede ser idéntica a la acción de sobresaltar a Ramírez. Que ésta es la propuesta de Feinberg se hace más claro en la distinción que él establece entre actos simples y causalmente complejos. Los actos simples son aquellos que no requieren que hagamos algo más (hemos estado llamando a éstas acciones primitivas); los actos causalmente complejos, tales como abrir o cerrar una puerta, sobresaltar o matar a alguien, requieren que hagamos, primero, como un medio, *algo más*.<sup>12</sup> Así, Feinberg dice: «A fin de abrir una puerta, debemos primero hacer algo más que *causará* que la puerta se abra; pero para mover un dedo propio, uno simplemente lo mueve —no se requiere ninguna actividad causal previa».<sup>13</sup> Él habla también de «sucesiones de actos causalmente conectados».

La idea de que abrir una puerta requiere de una actividad causal previa, un movimiento que causó que la puerta se abriera, no es sólo de Feinberg. Él cita a Austin con el mismo espíritu: «Puede hacerse que un único término descriptivo de lo que hizo cubra tanto una extensión menor como una mayor de sucesos, y luego se denominarán “consecuencias”, “resultados”, “efectos” o cosas similares, de este acto, los que queden excluidos por la descripción más estrecha.»<sup>14</sup> Arthur Danto ha trazado, en varios artículos, la distinción entre «actos básicos», tales como mover una mano, y otros actos que causan los actos básicos, tales como mover una piedra.<sup>15</sup>

11. Feinberg, «Action and Responsibility», p. 146. A mí me preocupa un asunto que no es central en el excelente artículo de Feinberg. Aun si mis advertencias están justificadas, su tesis no se ve seriamente afectada.

12. *Ibid.*, p. 145.

13. *Ibid.*, p. 147.

14. Austin, «A Plea for Excuses», p. 145.

15. Arthur Danto, «What We Can Do», «Basic Actions», «Freedom and Forbearance». Chisholm apoya la distinción en «Freedom and Action», p. 39.

Me parece que esta concepción de las acciones y de sus consecuencias contiene confusiones estrechamente relacionadas, pero muy básicas. Es un error pensar que cuando yo cierro la puerta por mi propia libre voluntad, *alguien* normalmente causa que yo lo haga, aun cuando sea yo mismo, o que cualquier acción mía, anterior o no, causa que yo cierre la puerta. Así que el segundo error es confundir lo que causa mi acción de mover la mano —el cerrar la puerta— con algo totalmente diferente —mi acción de cerrar la puerta. Y el tercer error, que es impuesto por los otros, es suponer que cuando yo cierro la puerta al mover la mano, realizo dos acciones numéricamente distintas (como lo tendría que hacer si se necesitara una para causar la otra). En lo que queda de este escrito desarrollo estas propuestas.<sup>16</sup>

Hay más que una señal de conflicto entre dos ideas incompatibles en Austin y en Feinberg. Como lo señalamos antes, Feinberg muestra alguna inclinación a tratar el mover la propia mano y abrir la puerta (y sobresaltar a Ramírez, etc.) como una y la misma acción que, de alguna manera, se estira o se contrae; pero también dice cosas que parecen contradecir esto, especialmente cuando sostiene que uno primero debe hacer alguna otra cosa para causar que la puerta se abra, a fin de abrir la puerta. La misma tensión se deja sentir en el pronunciamiento de Austin, pues él habla de diferentes términos descriptivos de *lo que el hombre hizo* —aparentemente, una y la misma cosa—, pero los términos «cubren» extensiones mayores o menores de sucesos. Los sucesos que cubren diferentes extensiones no pueden ser idénticos.<sup>17</sup>

Pienso que hay dificultades insuperables que obstaculizan el camino para considerar numéricamente distintas estas diversas acciones, las acciones primitivas como mover una mano y las acciones para cuya descripción nos referimos a las consecuencias.

16. La tesis de Danto de que si cierro la puerta al mover la mano, mi acción de cerrar la puerta es causada por mi mover la mano, ha sido adecuadamente criticada por Myles Brand, «Danto on Basic Actions»; Frederick Stoutland, «Basic Actions and Causality»; Wilfrid Sellars, «Metaphysics and the Concept of a Person». Mi meta es más general. Quiero oponerme a cualquier tesis que implique que si hago A al hacer B, entonces mi hacer A y mi hacer B deben ser numéricamente distintos.

17. Hay mayor discusión de estos asuntos en los ensayos 6 y 8.

Es evidente que la relación que hay entre el que la reina mueva su mano de tal manera que vierta veneno en la oreja del rey y el que ella lo mate no puede ser la relación de causalidad de sucesos. Si lo fuese, tendríamos que decir que la reina se causó a sí misma matar al rey. Esto no es lo mismo que decir que la reina produjo o hizo que sucediese que ella matase al rey; estas locuciones, aun cuando sean forzadas, no parecen claramente erróneas, pues no es claro que ellas signifiquen algo más que el que la reina se obligó a matar al rey. Pero, entonces, las locuciones no pueden ser causales en el sentido requerido. Pues supóngase que al mover la mano, la reina se causó a sí misma matar al rey. Entonces podríamos preguntar cómo hizo esta causación. La única respuesta que puedo imaginar es que lo hizo moviendo la mano de esa manera. Pero este movimiento, por sí mismo, fue bastante para causar la muerte del rey —no tenía objeto otra acción por parte de la reina. Ni hay ninguna razón (a menos que amplíemos la narración de una manera irrelevante) de por qué la reina habría querido causarse a sí misma matar al rey. Lo que quería hacer era matar al rey —esto es, hacer algo que causase su muerte. ¿No es absurdo suponer que, luego que la reina movió la mano de manera tal para causar la muerte del rey, le queda aún por realizar o completar algún hecho? Ella ha hecho su trabajo; sólo le queda al veneno realizar el suyo.

No será útil pensar en matar como una acción que comienza cuando se lleva a cabo el movimiento de la mano pero que concluye después, ya que, una vez más, cuando inquirimos por la relación entre estos sucesos, la respuesta debe ser que matar consiste en el movimiento de la mano y en una de sus consecuencias. Podemos reunirlos de esta manera porque el movimiento de la mano causó la muerte. Pero, entonces, al mover la mano, la reina estaba haciendo algo que causó la muerte del rey. Éstas son dos descripciones del mismo suceso —la reina movió la mano de esa manera; ella hizo algo que causó la muerte del rey (o para formularlo, como me agradaría, en términos de una descripción definida: el mover la reina la mano en esa ocasión fue idéntico a su hacer algo que causó la muerte del rey). Hacer algo que cause una muerte es idéntico a causar una muerte. Pero no

hay distinción alguna entre causar la muerte de una persona y matarla.<sup>18</sup>

Se sigue que lo que pensamos que era un suceso más atenuado —el matar— no llevó más tiempo que el movimiento de la mano, y no difirió de él.

La idea de que, en las circunstancias supuestas, matar a una persona difiere de mover de cierta manera la propia mano surge de una confusión entre un rasgo de la descripción de un suceso y un rasgo del suceso mismo. El error consiste en pensar que cuando se hace que la descripción de un suceso incluya una referencia a una consecuencia, entonces la consecuencia misma está incluida en el suceso descrito. El acordeón, que sigue siendo el mismo a través del comprimir y el estirar, es la acción; los cambios están en aspectos descritos o en descripciones del suceso. De hecho, hay muchas tonadas que se pueden tocar en el acordeón. Podríamos comenzar con «La reina movió la mano» y jalarlo a la derecha, añadiendo «causando así que se vaciara la ampolleta en la oreja del rey»; y ahora, un nuevo jalón, «causando así que el veneno entrara en el cuerpo del rey»; y, finalmente (si con esto tenemos bastante —pues no hay límites claros para las posibilidades de expansión), «causando así que el rey muriera». La expresión puede acortarse de varias maneras en los componentes del centro, de la izquierda o de la derecha o en cualquier combinación. Algunos ejemplos: «La reina movió la mano causando así la muerte del rey» (los dos extremos); o: «La reina mató al rey» (colapso a la derecha); o: «La reina vació la ampolleta en la oreja del rey» (el centro). También hay otra forma de mover el instrumento: podríamos *comenzar* con «La reina mató al rey», añadiendo «al verter veneno en su oído» y así sucesivamente —aumento a la izquierda. Muchas de estas expresiones son equivalentes; por ejemplo: «La reina mató al rey al verter veneno en su oreja» y «La reina vertió veneno en la oreja del rey, causando así su muerte». Y, obviamente, las descripciones más largas implicarán formalmente muchas de las más cortas.

18. Véase la nota 10. El argumento funciona si la aseveración de esta oración se debilita añadiendo «en el caso en el que una persona sea matada al hacer algo que cause su muerte».

Pero esta cantidad de descripciones relacionadas corresponde a un único *descriptum* —ésta es la conclusión en la que convergen todas nuestras consideraciones.<sup>19</sup> Cuando inferimos que él paró su auto del hecho de que al presionar un pedal un hombre causa que su automóvil pare, no transferimos la actuación de un suceso a otro ni inferimos que el hombre fue agente no sólo de una acción sino de dos. Podemos, ciertamente, extender la responsabilidad o la imputabilidad por una acción a la responsabilidad o a la imputabilidad por sus consecuencias, pero esto no lo hacemos achacándole al agente una nueva acción, sino señalando que su acción original tuvo esos resultados.

Debemos concluir, quizás no sin un sobresalto, que nuestras acciones primitivas, las que no hacemos al hacer alguna otra cosa, son meros movimientos del cuerpo —éstas son todas las acciones que hay. Nunca hacemos más que mover nuestros cuerpos; lo demás, se lo dejamos a la naturaleza.<sup>20</sup>

Puede parecer que esta doctrina, aun cuando no es tan mala como la vieja y mala doctrina de que lo único que hacemos es querer [*will*] que sucedan las cosas o prepararnos para actuar, comparte algunas de sus desventajas. Permítaseme indicar, brevemente, porqué pienso que esto no es así.

Primero, se dirá que algunas acciones requieren que hagamos otras a fin de producir aquéllas y, así, no pueden ser primitivas; por ejemplo, antes de darle al blanco, debo cargar y alzar mi pistola, luego apuntar y jalar el gatillo. Claro está que no niego que debo preparar el camino para algunas acciones realizando otras. La crítica vale sólo si muestra que algunas acciones no son primitivas. En este ejemplo, el desafío es demostrar que darle al blanco es una acción primitiva. Y ciertamente lo es, conforme al argumento que he dado, pues darle al blanco no es más que hacer algo que cause que se le

19. Esta conclusión no es nueva. La formuló claramente G.E.M. Anscombe, *Intention*, pp. 37–47. [Hay traducción al español: *Intención*, trad. Ana Isabel Stellino, Instituto de Investigaciones Filosóficas-UNAM/Paidós, México/Barcelona, 1991.] Yo seguí su ejemplo en el ensayo 1.

20. Aun cuando no es falsa, esta oración, sacada de contexto, es engañosa. Si muevo la tierra, esto *suena* como algo más que mover mi cuerpo. El argumento muestra que no es así. [Nota añadida en 1979.]



dé al blanco y esto, dadas las condiciones adecuadas, incluyendo un arma en la mano, lo puedo hacer sosteniendo mi arma en cierta posición y moviendo mi dedo en el gatillo.

Segundo, se dice con frecuencia que las acciones primitivas se distinguen por el hecho de que sabemos, quizás sin necesidad de observación o de evidencia, que estamos realizándolas, mientras que éste no es un rasgo de otros sucesos tales como darle al blanco. Pero es claro que podemos saber que cierto suceso está llevándose a cabo cuando se lo describe de una manera y no saber que se está llevando a cabo cuando se lo describe de otra. Incluso cuando estamos haciendo algo intencionalmente podemos no saber que lo estamos haciendo; es aún más claramente posible que esto sea verdadero acerca de acciones cuando éstas se describen en términos de sus resultados no buscados.

Finalmente, puede parecer una dificultad que las acciones primitivas no den cabida al concepto de intentar, pues las acciones primitivas son acciones que simplemente hacemos —nada puede obstaculizarlas, por así decir. Pero, ciertamente, dirá el crítico, hay algunas cosas que debemos tratar de hacer (como darle al blanco). Una vez más, sirve el mismo tipo de respuesta. Tratar de hacer una cosa puede simplemente ser hacer otra. Yo trato de prender la luz al presionar el interruptor. Pero simplemente presiono el interruptor. O, quizás, en alguna ocasión, incluso eso es un intento. Sin embargo, el intento consiste en algo que puedo hacer sin tratar de hacerlo; tan sólo muevo la mano, quizás.

El mismo hecho subyace en las dos últimas respuestas: haberlo intentado y saber que acontece no son características de los sucesos, sino de los sucesos en tanto que se los describe o se los concibe de una u otra manera. También es este hecho el que explica por qué podemos estar limitados, en nuestras acciones, a meros movimientos de nuestros cuerpos y, sin embargo, ser capaces, para bien o para mal, de construir presas, contener inundaciones, matarnos unos a otros o, de tiempo en tiempo, darle al blanco.

Podemos ahora volver a la pregunta acerca de la relación entre un agente y su acción. La conclusión negativa que habíamos alcanzado es ésta: la noción de causa nada tiene que ver, directamente, con esta relación. El conocimiento de que una acción *a* tiene cierto resultado nos permite describir al agente como la causa de ese resultado, pero

ésta es, tan sólo, una manera conveniente de redescubrir *a*, y de *ella*, como hemos visto, no tiene objeto decir que él es la causa. La causalidad nos permite redescubrir las acciones de formas en las que no podemos redescubrir otros sucesos; este hecho es una marca de las acciones, pero no produce ningún análisis de la actuación.

Decir que todas las acciones son acciones primitivas es tan sólo reconocer, quizás de manera engañosa, el hecho de que el concepto de ser primitivo, como el concepto de ser intencional, es intensional y, así, no puede marcar una *clase* de acciones. Si un suceso es una acción, entonces bajo alguna(s) descripción(es) es primitivo, y bajo alguna(s) descripción(es) es intencional. Esto explica por qué fracasamos en el intento de suponer que un concepto básico de actuación se aplica a las acciones primitivas, y de tenderlo a otras acciones definidas en términos de las consecuencias de las acciones primitivas: el intento fracasó porque no hay otras acciones, sólo hay otras descripciones.

El colapso de todas las acciones en las acciones primitivas, el cual se señala en la sintaxis por el efecto de acordeón, conduce a una gran simplificación del problema de la actuación, pues muestra que hay una relación entre una persona y un suceso, cuando es su acción, que es independiente de cómo se describan los términos de la relación. Por otra parte, no hemos descubierto ningún análisis de esta relación que no apele al concepto de intención. Averiguar si la intención puede entenderse en términos de ideas más básicas o simples es un problema que no he intentado tocar en este escrito.

## 4. LIBERTAD PARA ACTUAR

La tesis de que las acciones libres están causadas por estados y episodios como deseos, creencias, recuerdos y los impulsos de la pasión, recibe ataques de dos frentes. Tenemos las andanadas de quienes creen que pueden ver, o incluso probar, que la libertad es inconsistente con el supuesto de que las acciones están determinadas causalmente, al menos si las causas pueden rastrearse en sucesos fuera del agente. No me preocuparán directamente tales argumentos, puesto que no conozco ninguno que no sea más que superficialmente plausible. Hobbes, Locke, Hume, Moore, Schlick, Ayer, Stevenson y muchos otros han hecho lo que puede hacerse, o lo que jamás podría haberse necesitado, para eliminar las confusiones que puedan hacer que parezca que el determinismo frustra la libertad.

El otro ataque es más interesante. No se dirige en contra del determinismo en cuanto tal, sino en contra de la teoría causal de la acción. Si una acción libre es una que está causada de ciertas formas, entonces la libertad para actuar debe ser un *poder causal* del actor que interviene cuando se satisfacen ciertas condiciones. El defensor de la teoría causal no puede evadir, o al menos así lo parece, el reto de presentar un análisis de la libertad para actuar que la muestre como un poder causal. Aquí, el teórico causal se ha visto forzado a pasar a una posición defensiva, pues ninguna de las versiones propuestas responde todas las objeciones. Y lo que es peor, las fallas en los análisis conocidos muestran patrones recurrentes, lo que sugiere la imposibilidad de dar una solución satisfactoria. En un breve esbozo presento una dificultad conocida y molesta.

Es natural decir que una persona puede hacer algo (o que es libre de hacerlo) si todo lo que se requiere, si lo ha de hacer, es que *tenga*

*la voluntad [will]* de hacerlo. Pero «Lo haría si tuviese la voluntad de hacerlo», ¿expresa una disposición causal? Si tener la voluntad es un acto distinto de hacer, podría ser una causa, pero entonces surgiría la cuestión acerca de cuándo es libre un agente para tener la voluntad. Si tener la voluntad no es un acto distinto de hacer, entonces no puede ser una causa de hacer. El dilema no se resuelve sustituyendo querer por elegir, decidir, tener la intención o intentar; la mayoría de estas alternativas tan sólo hacen más obvio en qué extremo del dilema se clava la propuesta. Si no puede superarse esta dificultad, ¿no hemos de abandonar la teoría causal de la acción, aun cuando no la derrote un ataque frontal?

Creo que el dilema particular que acabo de mencionar puede resolverse. También pueden evadirse muchas otras objeciones a un análisis causal de la libertad. Sin embargo, queda una dificultad central que frustra permanentemente la formulación de un análisis satisfactorio, pero este hecho no debe persuadirnos de que la libertad para actuar no es un poder causal. Bosquejada de manera amplia, ésta es la tesis a favor de la cual argüiré ahora. Primero debe decirse algo más acerca de la idea de un poder causal.

Por un poder causal me refiero a una propiedad tal de un objeto que un cambio de cierto tipo en el objeto causa un suceso de otro tipo. Claro está que esta caracterización requiere aclararse mucho. Por ejemplo, apelar a tipos de cambio invita a hacer un examen de leyes causales, contrafácticos y clases naturales. No obstante, espero evitar los peores problemas que rondan por aquí, reduciendo los supuestos. En particular, no consideraré como dado que los poderes causales puedan analizarse mediante condicionales subjuntivos, aun cuando gran parte del examen recaerá en lo adecuado de análisis propuestos en términos de condicionales subjuntivos.

El concepto de poder causal es indiferente a la distinción intuitiva entre lo activo y lo pasivo. Podemos pensar en la solubilidad como en una característica pasiva y en la de ser solvente como en una activa. Sin embargo, son lógicamente equivalentes al ser la primera una propiedad de las cosas que, por algunos cambios en ellas, causa que se disuelvan, y la última, una propiedad de las cosas que, por algunos cambios en ellas, causa que otras cosas se disuelvan. Ambas son igualmente poderes causales.

Muchos de los ataques en contra de la teoría causal de la acción libre se apegan a poderes como la solubilidad, como si sólo tales poderes pudiesen proporcionar análogos de ser libre para actuar. Esto fuerza a la teoría de una manera injustificada. Si una sustancia es soluble, ciertas eventualidades causarán que se disuelva; la analogía sugiere que una acción libre es una en la que algo que le sucede al agente causa que éste actúe. Sin embargo, una analogía igualmente buena diría que una acción libre es una en la que un cambio en el agente causa que algo suceda fuera de él. Éste, de hecho, sería el significado natural de «poder causal» atribuido a una persona; ella tiene el poder de causar que sucedan cosas. De nuevo, sin embargo, están equivocados los detalles del paralelismo. Es verdad que una persona es libre de producir lo que sea que producirían las acciones que es libre de realizar. Pero este concepto de lo que un agente es libre de producir depende de una idea previa acerca de lo que es libre de hacer. Un hombre podría tener el poder de destruir el mundo, en el sentido de que si simplemente presionase un botón ante él se destruiría el mundo. Sin embargo, él no sería libre de destruir el mundo a menos que fuese libre de presionar el botón.

Otro tipo de poder causal consiste en que se causa que se dé cierto cambio en un objeto que lo posee si un cambio previo tiene lugar en el objeto. Un ejemplo podría ser la propiedad de un tanque que se autosella. Si el tanque empieza a escurrir, esto causa que se selle el escurrimiento. Propiedades similares son: estar en equilibrio, estar inestable, ser homeostático. Si la libertad para actuar es un poder causal, pertenece a esta categoría.

Hay una pequeña diferencia, quizás, entre «El calor causó que la flor se marchitase» y «El calor causó el marchitamiento de la flor». Pero puede haber una gran diferencia entre «El calor causó que Samantha regresara a Patna» y «El calor causó el retorno de Samantha a Patna». La primera implica o sugiere fuertemente una limitación en la libertad de acción de Samantha; no así la segunda. Al poner a prueba la plausibilidad de la aseveración de que la libertad para actuar es un poder causal debemos, por tanto, tener en mente la lectura neutral. Una *acción* puede ser causada sin que se cause al *agente* a que la realice. Aun cuando la causa es interna, tenemos la sensación de una diferencia. «El deseo causó que lo

hiciera» sugiere una falta de control que podría presentarse como excusa de la acción porque hace el acto menos voluntario; «El deseo fue la causa de que lo hiciera» deja abierta la cuestión de la libertad.

Dejando ahora estas cuestiones preliminares, consideremos argumentos ideados para mostrar que ser libre para realizar una acción no puede ser un poder causal. Podemos comenzar con algunas de las aseveraciones en el artículo de Austin «Ifs and Cans». Conforme a Austin, G.E. Moore sugirió las siguientes tres proposiciones, las dos primeras de manera explícita y la tercera implícitamente:

- (1) «Podría haber» significa simplemente «Podría haber si lo hubiese elegido».
- (2) Por «podría haber si lo hubiese elegido» podemos sustituir «debería haber sido si lo hubiese elegido».
- (3) Las cláusulas «si» en estas expresiones enuncian las condiciones causales conforme a las que se habría seguido que yo podría o debería haber hecho una cosa diferente de lo que realmente hice.

Austin arguye que (1)–(3) son falsas. Yo argüiré que, a pesar de todo lo que Austin dice, podrían haber sido verdaderas (aun cuando hay un aspecto importante acerca del que Austin tiene razón: uno de los disyuntos de (3) es falso).

Supongamos que la tesis básica de Moore era que decir que un agente podría haber hecho algo o puede ahora hacer algo (esto es, que era o que es libre para hacerlo) significa que, si hubiese elegido hacerlo, lo habría hecho (o que si ahora elige hacerlo, él lo hace o lo hará). He transpuesto la tesis de la primera a la tercera persona a fin de anticiparme a escrúpulos no pertinentes acerca de realizaciones y he generalizado un tanto los tiempos verbales (quedan dificultades sobre el tiempo y los tiempos verbales, pero no creo que tengamos que resolverlas aquí). Para quedarnos con una formulación, consideremos que la tesis de Moore es: «A puede \_\_\_\_\_» significa «Si A elige \_\_\_\_\_ (inf.), entonces él \_\_\_\_\_ (3a. p.)» (un ejemplo de esto sería: «Ramírez puede anclar su bote en el canal» significa «Si Ramírez elige anclar su bote en el canal, entonces él ancla su bote

en el canal».<sup>1</sup> Examinaré el ataque de Austin en contra de Moore a la luz de esta propuesta.

Si aceptamos la sugerencia, entonces «A puede \_\_\_\_» sí significa lo mismo que (o, de cualquier manera, es lógicamente equivalente a) «A puede \_\_\_\_ si lo elige». Pues dada la sugerencia de Moore, «A puede \_\_\_\_» significa «A \_\_\_\_ (3a. p.) si lo elige», mientras que «A puede \_\_\_\_ si lo elige», se expande en «A \_\_\_\_ (3a. p.) si lo elige, si lo elige» y estas dos últimas oraciones son lógicamente equivalentes.

Moore no ve ninguna diferencia importante entre «A \_\_\_\_ (3a. p.) si lo elige» y «A puede \_\_\_\_ si lo elige». Y dado su análisis sugerido de «puede», de seguro tiene razón, pues la segunda oración se convierte, como lo acabamos de ver, en «A \_\_\_\_ (3a. p.) si lo elige, si lo elige», que es lógicamente equivalente a «A \_\_\_\_ (3a. p.) si lo elige». Austin arguye en contra de esta equivalencia, pero el argumento es defectuoso. Él asume como obvio que si «A \_\_\_\_ (3a. p.) si lo elige» significa lo mismo que «A puede \_\_\_\_ si lo elige», entonces «\_\_\_\_ (3a. p.)» y «puede \_\_\_\_» deben significar lo mismo, y esto es patentemente falso. Pero el supuesto está equivocado. «\_\_\_\_ (3a. p.)» y «puede \_\_\_\_» no significan lo mismo y, sin embargo, son equivalentes las oraciones «A \_\_\_\_ (3a. p.) si lo elige» y «A puede \_\_\_\_ si lo elige», concediendo la sugerencia de Moore de que «puede \_\_\_\_» significa «\_\_\_\_ (3a. p.) si lo elige».<sup>2</sup>

Austin subraya el hecho de que no todas las cláusulas «si» introducen condiciones causales; algunas cláusulas en las que «si» es la conectiva principal son condicionales causales y algunas no. Austin obviamente tiene razón al decir que «A puede \_\_\_\_ si lo elige» no es un condicional causal, al menos conforme al análisis que da Moore de «puede». La razón es que el mismo «A puede \_\_\_\_» (estamos

1. Las indicaciones que figuran después de algunas de las líneas que señalan espacios son las siguientes: «(inf.)», poner el verbo de acción en infinitivo; «(3a. p.)», poner el verbo en 3a. persona del singular (presente de indicativo); y «(ría)», poner el verbo en 3a. persona del singular (futuro hipotético o condicional). [N. del T.]

2. En «Ifs, Cans and Causes» Keith Lehrer señala la equivalencia, en el análisis de Moore, de «Yo haré si lo elijo» y «Yo puedo si lo elijo». Pero luego copia el *non sequitur* de Austin: «Yo puedo, si lo elijo» es claramente no equivalente a «Yo haré, si lo elijo» porque «Yo puedo» no es equivalente a «Yo haré».

suponiendo) es un condicional, cuyo antecedente es «Él elige». Así que «A puede \_\_\_\_ si lo elige» implica formalmente el consecuente, lo que muestra que la oración más larga no puede ser un condicional causal. Sin embargo, sería torpe razonar que puesto que «A puede \_\_\_\_ si lo elige» es lógicamente equivalente a «A \_\_\_\_ (3a. p.) si lo elige» y el primero no es un condicional causal, entonces el último tampoco lo es.

Austin parece haber pensado que la palabra «si» tiene en los condicionales causales un significado diferente que en otros lugares, pero la evidencia a favor de esto no es muy convincente. No hay buenas razones para decir que «si» tiene significados diferentes en «Si un número es divisible entre cuatro, es divisible entre dos» y en «Si comes bastantes semillas de manzana, tendrás envenenamiento de arsénico», aun cuando, presumiblemente, el último es un condicional causal y no así el primero. La diferencia no es semántica, sino epistemológica; sostenemos que la primera oración es verdadera por razones lógicas o aritméticas; sostendríamos que la segunda oración es verdadera si pensásemos que comer semillas de manzana *causase* envenenamiento de arsénico.

La cuestión del significado es pertinente para la siguiente propuesta. Austin dice que es característico de los condicionales causales que la contraposición es una forma válida de inferencia. Esto es seguramente correcto y podemos ver que la contraposición funciona para «A \_\_\_\_ (3a. p.) si lo elige». Pero, ¿hay algunas oraciones «si» para las que no funcione la contraposición? (Si las hay, deberemos admitir que «si» es ambiguo.) Pienso que no. No estamos inclinados a *razonar* de «A puede \_\_\_\_ si lo elige» a su contrapositiva, pero pienso que la explicación no es que la implicación formal falle. Una mejor explicación es ésta: conforme al análisis de Moore, «A puede \_\_\_\_ si lo elige» significa «A \_\_\_\_ (3a. p.) si lo elige, si lo elige», la que a su vez resulta ser lógicamente equivalente a su consecuente. La cuestión acerca de si «A puede \_\_\_\_ si lo elige» implica formalmente su contrapositiva es, por tanto, tan sólo la cuestión acerca de si «A puede \_\_\_\_» implica formalmente «Si A no puede \_\_\_\_, entonces no elige \_\_\_\_». Claro está que lo implica formalmente, de la manera formal y vacua en la que una contradicción implica formalmente cualquier cosa que sea. Éste no



es el tipo de implicación formal que invite a un razonamiento serio fuera de la lógica. El argumento de Austin nos ayuda a ver, una vez más, que «A puede \_\_\_\_ si lo elige» no es un condicional causal, pero no muestra que un «si» causal esté en perspectiva—esto es, en el «puede».<sup>3</sup>

En una reseña de los artículos de Austin,<sup>4</sup> Roderick Chisholm llega, por una ruta algo distinta, esencialmente a la misma conclusión que yo he alcanzado, de que Austin no ha dado ninguna razón concluyente para rechazar el análisis de Moore de «puede» en términos de «lo hace si lo elige», ni se ha mostrado que las oraciones resultantes no sean condicionales causales. Chisholm cree que Austin ha mostrado que «Él puede \_\_\_\_ si lo elige» no es un condicional y que hay dificultades acerca de «elige» como el verbo apropiado para el antecedente. Yo no pienso que Austin logre demostrar ninguna de estas cosas, pero esto no importa con respecto al argumento de Chisholm, pues el argumento de Chisholm, si es bueno, no sólo se aplica a «elige», sino también a muchos otros verbos que podríamos poner en su lugar, como «intenta».

La propuesta de Chisholm es que mientras que podría ser verdadero que si un hombre fuese a elegir o intentar realizar una acción, entonces él realizaría esa acción, sin embargo él podría no ser capaz de elegir o de intentar, en cuyo caso él no podría realizar la acción. Así que podría ser verdad que una persona \_\_\_\_ (ría) si lo intentase y falso, sin embargo, que pudiese \_\_\_\_.

Pienso que el argumento de Chisholm es muy importante y que tiene razón al decir que pasarlo por alto es cometer un «error de primera calidad, básico».<sup>5</sup> Lo que el argumento muestra es que el antecedente de un condicional causal que intente analizar «puede», «pudiese» o «ser libre para» no debe contener, como su verbo do-

3. Por razones análogas yo sostendría que «Hay bizcochos en el aparador si los quieres» contiene un «si» normal e implica formalmente su contrapositiva. Dejamos de ver esto porque suponemos que el mero querer no puede producir los bizcochos; por tanto, «Hay bizcochos en el aparador si los quieres» es verdadera sólo si hay bizcochos en el aparador. Este supuesto contradice el antecedente de la contrapositiva, así que en lugar de inferir la contrapositiva, simplemente nos desconcierta la pregunta acerca de si es verdadera.

4. «J.L. Austin's *Philosophical Papers*», pp. 20-25.

5. *The Refutation of Determinism*, p. 145.

minante, un verbo de acción o cualquier verbo que haga que tenga sentido la pregunta ¿puede alguien hacer *eso*?

Si no estoy equivocado, M.R. Ayers ha supuesto que porque el antecedente de un condicional causal que analiza la libertad para actuar no puede estar dominado por un verbo de acción, ningún análisis causal es posible. Su argumento parece ser que en todos los otros casos de condicionales causales, el antecedente es algo que podemos hacer verdadero de manera independiente del efecto supuesto. Pero en el caso de la acción, nos dice: «que poner a prueba la habilidad propia para hacer algo no es lo mismo que poner a prueba la verdad de una hipótesis, se sigue simplemente del truismo de que a fin de hacer algo no siempre es necesario hacer alguna otra cosa antes». El autor concluye: «Así queda excluida la posibilidad de cualquier análisis hipotético del poder personal.»

La conclusión no se sigue de la premisa. Lo que se sigue del truismo de que «a fin de hacer algo no siempre es necesario hacer otra cosa antes» es sólo que si un condicional causal ha de analizar la libertad de actuar, el antecedente no puede tener como su verbo principal un verbo de acción. Claro está que esto no significa que el agente mismo no pueda, en general, poner a prueba la verdad del condicional haciendo alguna otra cosa —haciendo verdadero tal antecedente. Sin embargo, no se sigue que él no pueda poner a prueba de alguna otra manera si él tiene la capacidad de hacer algo, ni se sigue que la capacidad pueda no consistir en un poder causal. Todo lo que podemos concluir es que la puesta a prueba no puede llevarse a cabo (siempre, de cualquier manera) con el agente haciendo algo para crear las condiciones para poner en acto el poder.

Keith Lehrer ha insistido, sobre bases mucho más generales, que ningún condicional causal puede dar el significado de «A es libre para \_\_\_\_ (o puede \_\_\_\_)». Él sostiene que «A \_\_\_\_ (3a. p.) si se da la condición C» es consistente con los dos enunciados siguientes: «A no puede \_\_\_\_ si no se da la condición C» y «C no se da». Sin embargo, si estas dos últimas oraciones *son* verdaderas, A no puede \_\_\_\_\_. En palabras de Lehrer, «es lógicamente posible que alguna condición, que es una condición suficiente para causar que

una persona haga algo, deba ser también una condición necesaria para que sea capaz de hacerlo, y que esa condición no suceda».<sup>6</sup>

Aun cuando Lehrer no parece haberlo notado, su argumento nada tiene que ver en particular con la acción. Si su línea de razonamiento es correcta, muestra que ninguna atribución de un poder o de una disposición jamás es equivalente a un condicional (sea que el condicional se interprete como ley causal o como subjuntivo). Así, conforme al argumento de Lehrer, decir que algo es soluble en agua no puede significar que se disuelve si se sitúa en el agua, puesto que puede no ser soluble y, sin embargo, situarlo en el agua lo haría disolverse.<sup>7</sup> Hay diversas maneras como uno podría intentar enfrentar esta propuesta. Pero, para mis propósitos, bastará señalar que incluso si el argumento de Lehrer mostrase que ninguna disposición o poder se analiza correctamente mediante un condicional subjuntivo, esto no socavaría la aseveración de que ser capaz de hacer algo es tener un poder causal.

Claro está que hay casos en los que querer, tratar, elegir hacer algo o incluso tener la intención de hacerlo nos impide hacer esa misma cosa, aun cuando pueda haber algún sentido en el que podríamos haberla hecho (si no hubiésemos querido, tratado, etc.). Los moralistas, desde Aristóteles hasta Mill, han señalado que es improbable que tratar de ser feliz produzca la felicidad y Schlick estaba tan convencido de esto que revisó el hedonismo para que no se expresara como «haz lo que puedas para ser feliz», sino como «preparate para la felicidad». Austin nos da un ejemplo más pedestre: un hombre puede tratar (elegir, querer, tener la intención de) de meter un *putt* [en el golf] y fallar, sin admitir que no podía. Ninguno de estos casos es como el de Lehrer. Más bien, son casos en los que uno puede y, sin embargo, aun cuando uno trate (elija, tenga la intención de, quiera), no lo logra.

6. «An Empirical Disproof of Determinism», p. 196.

7. Le agradezco a G.E.M. Anscombe por corregir mi interpretación del argumento de Lehrer. La respuesta correcta a Lehrer es simplemente que si uno analiza solubilidad mediante un condicional causal, no se puede conceder, de manera consistente, que lo que causa la disolución es también una condición necesaria de la solubilidad, puesto que en ese caso las únicas cosas solubles estarían disueltas. Algo similar sucede con los análisis condicionales de la libertad. [Nota añadida en 1977.]

Chisholm propone una forma de eludir esta dificultad. Él sugiere que *A* puede hacer algo siempre que haya algo (diferente o igual) tal que *A* trate (elija, quiera, tenga la intención de) de hacer *eso*, luego él hace la primera cosa. En el caso del *putt*, esto significa que el agente puede meter el *putt* si hay alguna otra cosa, quizás golpear la pelota una pulgada a la izquierda del hoyo, tal que si él intenta hacer *eso*, entonces meterá el *putt*.

Ni Chisholm ni yo pensamos que éste sea el análisis correcto, sólo que responde la objeción de Austin a la propuesta de Moore. Pero hay otro hecho (no señalado por Chisholm) que apoya la revisión que hace Chisholm de Moore. Cuando describimos nuestras acciones, incluimos no sólo lo que hacemos intencionalmente, sino también muchas cosas que hacemos no intencionalmente. Si se sostiene, como yo lo hago, que las acciones no intencionales son intencionales en otras descripciones, entonces la propuesta puede formularse diciendo que las descripciones de acción incluyen descripciones de acciones como intencionales y algunas otras descripciones de esos mismos sucesos. Por ejemplo, yo puedo intencionalmente depositar una moneda sobre el mostrador de un puesto de periódicos en París con la intención de pagar una compra. No intencionalmente pongo un dracma en lugar de un franco provocando, con ello, una interesante expresión en *argot*. Mi única acción puede caracterizarse correctamente como: depositar una moneda, depositar un dracma, provocar una interesante expresión en *argot*. Sólo en la primera caracterización mi acto es intencional.

Una manera importante de describir las acciones (y otros sucesos) es en términos de sus efectos o consecuencias. Provocar una expresión interesante en *argot* caracteriza mi acto de depositar la moneda en términos de una consecuencia, a saber, la respuesta provocada. La respuesta provocada, claro está, no es *parte* de mi acción, sólo parte de su descripción. Podemos llamar al efecto causado por mi acción la *terminación* de la acción. Mis acciones incluyen, entonces, lo que hago intencionalmente y cualquier cosa que yo haga cuya terminación la cause una acción. La segunda cláusula no añade nuevos sucesos, sólo nuevas formas de describir el viejo.<sup>8</sup>

8. Véase también el ensayo 3.

De manera que, si tengo razón, el análisis del concepto de una acción se presenta en dos etapas, la primera introduce directamente la intención y la segunda extiende el concepto a acciones en el primer sentido, redescritas en términos de consecuencias no buscadas o de otras características no buscadas. Puesto que lo que hacemos ciertamente está incluido en lo que podemos hacer, el patrón de análisis que se aplica al concepto de acción debe funcionar para la noción de lo que un agente puede o es libre de hacer. Es libre de hacer esas cosas que él haría si eligiese (quisiese, intentase) hacerlas; pero también es libre de hacer esas cosas cuyas terminaciones serían causadas por una acción que él es libre de hacer en el primer sentido.

Ésta es una manera muy generosa de enumerar las cosas que podemos hacer —incluye todas las cosas que podríamos producir mediante nuestra acción intencional, sea de manera planeada o por accidente, por una suerte ciega o por una maquinación maestra. En este amplio sentido de «poder» o de «podría», cada uno de nosotros podría hacer un millón de dólares en el mercado de cambios, casarse con una estrella de cine o, incluso, ponerle fin a la guerra. Si tan sólo supiéramos cómo.

Para los propósitos prácticos, con frecuencia estamos interesados en conceptos más limitados de lo que podemos hacer —lo que realmente hacemos si queremos, intentamos o tratamos de hacer *eso*, por ejemplo, o lo que podemos hacer confiablemente. Pero el concepto más amplio que he introducido es básico, al menos en el sentido de que otros podrían definirse en términos de él —suponiendo, claro está, que hasta aquí nuestro análisis es correcto. Es, simplemente, un concepto más inclusivo que el que usualmente empleamos, porque incluye, entre las cosas que podemos hacer, las cosas que haríamos sólo por azar o si supiésemos más de lo que sabemos. Pero no es tan amplio como para incluir lo meramente posible, lógica o físicamente —lo que alguien podría hacer sin que la naturaleza se lo impidiera. Esta categoría podría abarcar cosas que yo no podría hacer intencionalmente y que no serían causadas por ningún acto intencional que yo pudiera realizar, como correr una milla en cuatro minutos.

Ahora nos hemos decidido por una teoría de dos partes acerca de lo que un hombre puede hacer: una parte se dirige a explicar lo que él puede hacer intencionalmente y la otra parte extiende esto a lo

que lo intencional causaría. Deseo concentrarme en la primera parte y, por el momento, en el verbo principal del antecedente. Decir que un hombre puede hacer intencionalmente *x* es decir que él lo haría si —¿qué?

El problema es el que esbozamos al principio. A fin de ser elegible como una causa, el suceso mencionado debe estar separado de la acción; pero si está separado de la acción, parece que hay siempre la posibilidad de preguntar, acerca de él, si el agente tiene libertad de hacerlo. La objeción se aplica a elegir, tratar, tener la voluntad o la intención. Ninguna de éstas es, plausiblemente, la causa de una acción porque, normalmente, éstas son formas de caracterizar la acción elegida, querida, pretendida o intentada, no son descripciones de otras acciones, sucesos o estados. Así, si un hombre trata de batear un *home run* y lo logra, su intento *es* su éxito y no puede ser su causa. En ocasiones, elegir o tener la voluntad de hacer algo son actos mentales realizados antes de hacer las cosas pero, en estos casos, surge la cuestión acerca de si el agente es libre para hacerlos. El análisis falla por otra razón, pues un hombre puede elegir o decidir el lunes tragarse un botón el martes y, sin embargo, no hacerlo, y esto no necesariamente mostraría que él no podía tragarse un botón o el lunes o el martes —o siquiera que él no podía, el lunes, tragarse un botón el martes. Finalmente, hemos de recordar que una persona puede ser capaz de hacer algo y puede realmente hacerlo intencionalmente y, sin embargo, nunca haber elegido, decidido, intentado o pretendido hacerlo.

La única esperanza para el análisis causal es encontrar estados o sucesos que sean condiciones causales de acciones intencionales, pero que ellos mismos no sean acciones o sucesos acerca de los cuales pueda formularse inteligiblemente la cuestión acerca de si el agente puede realizarlas. Los más elegibles de tales estados o sucesos son las creencias y los deseos de un agente que *racionalizan* una acción, en el sentido de que sus expresiones proposicionales ponen la acción bajo una luz favorable, proporcionan una explicación de las razones que el agente tuvo para actuar y nos permiten reconstruir la intención con la que actuó. Por ejemplo, supongamos que un hombre asierra un piano a la mitad porque desea arrojar el piano por la ventana y cree que aserrar el piano a la mitad lo ayudará en su cometido. Su razón para aserrar el piano

a la mitad fue que deseaba arrojar el piano por la ventana —a lo que podemos añadir que él creía que podía hacer esto si primero aserraba el piano a la mitad. Obviamente sus razones, o las creencias y los deseos que les corresponden, explican por qué actuó como lo hizo. Y podemos ver que la intención por la que aserró el piano a la mitad era para arrojar el piano por la ventana.

Así, ahora sugiero que consideremos la siguiente formulación de la primera parte de nuestra teoría sobre qué puede hacer un hombre:

*A puede hacer x intencionalmente (en la descripción d) significa que si A tiene deseos y creencias que racionalizan x (en d), entonces A hace x.*

Varios problemas previos parecen resolverse por este análisis. La condición antecedente (*A* tiene deseos y creencias que racionalizan *x*) es anterior a la acción y está separada de ella, y, así, es adecuada para ser una causa (en este caso es un estado más bien que un suceso —pero esto podría cambiarse en este sentido: «llegar a tener deseos y creencias que racionalicen *x*»). La condición antecedente no menciona algo que sea una acción, así que no surge la pregunta acerca de si el agente puede hacerla. El meollo de la cuestión no es que los deseos y las creencias no estén jamás bajo el control del agente sino, más bien, que llegar a tenerlos no es algo que haga un agente. No quiero sugerir que la naturaleza de las creencias y de los deseos de un agente y la cuestión de cómo se adquieren no sean pertinentes con respecto a cuestiones acerca de cuán libre es él o lo son sus acciones. Pero estas cuestiones se encuentran en un nivel diferente y más complejo que las de nuestro estudio actual.<sup>9</sup>

9. No todos concuerdan en que ayuda a resolver nuestro problema el distinguir entre condiciones antecedentes que son acciones y condiciones antecedentes que no lo son. Así, Chisholm dice que los análisis condicionales de «puede hacer \_\_\_\_\_» son cuestionables incluso si el antecedente del *analysans* no concierne a una acción (*op. cit.*, p. 25). Por tanto, en conexión con la teoría que se considera, Chisholm diría que el agente no es libre de \_\_\_\_\_ aun si él haría \_\_\_\_\_ si tuviese los deseos y las creencias apropiadas, siempre que él no pudiese tener el deseo o la creencia. En este ensayo rechazo esta tesis. Sostengo que ha y un sentido básico en el cual somos libres de \_\_\_\_\_ (podemos \_\_\_\_\_) siempre que todo lo que se necesite para la acción (además de otras circunstancias presentes) sean las actitudes y las creencias correctas. La cuestión acerca de si podemos tener esas actitudes y creencias, a su

Añadiéndole «intencionalmente» al *definiendum* hemos respondido una de las objeciones de Austin contra un análisis condicional de «poder», pues una de sus objeciones a «Lo haré si lo elijo» era que el consecuente sólo implica formalmente «puedo». De manera similar, si analizásemos «A puede hacer *x*» como «Si A tiene actitudes que racionalizan *x*, entonces A hace *x*», Austin sin duda se quejaría de que «A hace *x*» nuevamente implica formalmente «A puede hacer *x*», mostrando que hay algo extraño acerca del condicional. La dificultad desaparece cuando el *definiendum* es «A puede hacer *x* intencionalmente», pues esto *no* se sigue de «A hace *x*». Pero, más importante, la razón para el cambio muestra que, en primer lugar, la objeción no era válida. De hecho, hacer *x* *no* implica formalmente que uno pueda hacer *x*.<sup>10</sup> La palabra «hacer» sugiere intención en algún lugar, pero en el presente contexto sirve sólo como un espacio para llenarlo por cualquier verbo cuyo sujeto pueda nombrar o describir una persona. Por el hecho de que hoy al mediodía cumplí exactamente 55 años, 72 días, 11 horas y 59 minutos de edad, no se sigue que yo tuve libertad para llegar a esta edad (más bien que a otra) en esta fecha o que yo podría hacerlo. Simplemente no estaba en mi poder hacerlo —ningún deseo o creencia de mi parte desempeñó algún papel causal en que yo llegara a esta edad en esta fecha, ni podría haberlo hecho. Si, por otra parte, actitudes racionalizantes causan una acción mía, entonces no sólo la acción sucede, sino que, en la descripción racionalizada, es intencional. Esto, claro está, es lo que debíamos esperar: lo que un agente hace intencionalmente es lo que está en libertad de hacer y tiene razones adecuadas para hacerlo.

La discusión de los últimos párrafos puede ayudarnos a poner en perspectiva otra perplejidad conocida. Es natural suponer que una acción que uno está en libertad de realizar es una acción que uno tam-

vez (excepto en circunstancias especiales), no es pertinente. Hay quienes sostienen que admitir esto es, al menos, conceder que «puede» tiene un significado diferente en «puede \_\_\_\_\_» (queriendo decir «es libre de \_\_\_\_\_») y en «puede creer . . .» o «puede querer . . .». Yo no veo que se siga esta conclusión, pero quizás sea verdad. [Nota añadida en 1979.]

10. David Pears hace la misma observación, aun cuando por razones algo diferentes, en «Ifs and Cans», p. 382.



bién está en libertad de *no* realizar. De manera similar, encontramos que muchos filósofos sostienen que una acción se realiza libremente sólo si el agente, cuando la realizó, tenía libertad de abstenerse; él podía haber actuado de otra manera. Pues supongamos que él no podía hacer otra cosa; ¿no significa esto que no tenía elección? Si la teoría causal nos lleva a esto, ¿por qué no tiene razón el libertario cuando niega que la libertad pueda reconciliarse con la teoría causal? O, para tomar el caso meramente potencial, ¿no es una pretensión vacía decir que un hombre es libre para realizar una acción si no tiene también libertad para no realizarla? Ciertamente, libertad significa la existencia de alternativas.

La dificultad (que recientemente puso de manifiesto Harry Frankfurt)<sup>11</sup> es que si decimos que un hombre es libre de hacer *x* sólo si su hacer *x* depende de si vale o no la condición de actitud (él elige hacer *x*, decide, lo quiere, tiene actitudes racionalizantes) entonces encontramos contraejemplos en casos de sobredeterminación. Lo que un hombre hace con su propia libre voluntad —una acción hecha por elección y con propósito, causada por sus propios deseos y creencias— puede ser algo que él podría haber sido causado a hacer de otra manera si faltaran la elección o los motivos.

Dos intuiciones parecen estar en pugna y el territorio amenazado por la destrucción lo ocupa la teoría causal. Las intuiciones son, por una parte, la tesis de que no podemos tener libertad de hacer lo que, de cualquier manera, estaríamos causalmente determinados a hacer y, por la otra, la sensación de que si elegimos hacer algo y lo hacemos porque lo elegimos, entonces la acción es libre, sin importar lo que habría sucedido si *no hubiésemos* elegido.

Resuelve la perplejidad un descubrimiento que hicimos en otro contexto. Lo que depende del agente es la realización intencional de un acto de cierto tipo. Es verdad que, en ocasiones, puede suceder que lo que un hombre hace intencionalmente, podría haber sido causado a hacerlo, de cualquier manera, por fuerzas extrañas. Pero en tal caso, lo que habríamos hecho no habría sido intencional. Así que incluso en los casos de sobredeterminación, algo recae en el agente. De hecho, no es *lo que él hace* (cuando se describe de una manera que deja abierto si acaso fue intencional) sino si acaso lo

11. «Alternate Possibilities and Moral Responsibility».

hace intencionalmente. Su acción, en el sentido en el que la acción depende de la intencionalidad, sucede o no, como él quiera; lo que él hace, en el sentido más amplio, puede suceder, sea que él lo quiera o no.

Pienso que estas reflexiones muestran que es un error suponer que añadimos algo al análisis de la libertad cuando decimos que un agente es libre de hacer algo si puede hacerlo *o no*, como él quiera (elija, etc.). Pues si esto significa que él actúa libremente sólo si no hubiera hecho lo que hizo si hubiesen estado ausentes las condiciones de actitud, es falso. Y si significa que él no habría actuado intencionalmente si las condiciones de actitud hubiesen estado ausentes, entonces esto es verdad, pero es gratuito decirlo, puesto que es una verdad lógica.

He estado examinando el concepto de lo que una persona puede hacer, o tiene libertad de hacer, cuando se interpreta este concepto como si incluyera lo que puede hacer, sea que lo sepa o no, haciendo algo intencional. Conforme a la propuesta actual, un hombre está en libertad de asesinar a un futuro tirano si está cerca del futuro tirano, tiene una pistola y de alguna manera la oportunidad, aun cuando no tenga razón para pensar que el hombre junto a él es el futuro tirano. Pues *si él creyese* que el hombre junto a él es el futuro tirano y él deseara, sobre todo, matar al futuro tirano, entonces, dadas las circunstancias propicias, él lo mataría. Obtenemos una versión más limitada de lo que él puede hacer si decimos que puede matar al futuro tirano sólo si, dado que él quiere matar al futuro tirano, él lo hace. Pues entonces, puesto que él no lo hará intencionalmente a menos que sepa o crea que el hombre junto a él es el futuro tirano, él no tiene libertad de hacerlo a menos que tenga el conocimiento. Obviamente este sentido más restringido de ser capaz de hacer algo es importante cuando se tiene que hacer responsable a un hombre por no actuar.

Mi propósito al señalar cómo podemos ajustar las condiciones en las que un hombre actúa para acomodar diversas interpretaciones de la libertad es hacer clara la medida en la que el problema de analizar la libertad para actuar se encuentra entramado con el problema de definir la acción intencional (y con esto, la acción en general, puesto que las acciones no intencionales son acciones intencionales en otras descripciones). La razón es obvia; decir cuándo

un agente tiene libertad de realizar intencionalmente (esto es, con cierta intención) una acción es enunciar condiciones en las cuales él realizaría la acción; explicar la realización de una acción con cierta intención es decir que las condiciones están satisfechas. Claro está que este estilo de análisis y de explicación funciona sólo si la satisfacción de las condiciones (que, estamos suponiendo, son causales) siempre conduce a la realización de la acción. Y, así, si podemos analizar o bien la libertad de actuar (en nuestro sentido amplio) o actuar con una intención, estaremos en posesión de una ley que diga que siempre que se satisfacen ciertas condiciones, un agente realizará una acción con cierta intención.

Puesto que, en la teoría causal, la acción intencional se define en parte por sus causas, quizás no debería sorprendernos que tan sólo con saber el análisis de la acción intencional estaríamos en el dominio de una ley de conducta. Empero, ¿cómo puede surgir lo que parece ser una generalización empírica a partir de consideraciones totalmente analíticas? En un momento volveré a esta cuestión.

Así que puede afirmarse que estas cosas se sostienen juntas: una ley que enuncia condiciones en las cuales los agentes realizan acciones intencionales, un análisis de la libertad de actuar que la hace un poder causal y un análisis causal de la acción intencional. Se sostienen juntas o caen juntas. En mi opinión, caen juntas si lo que queremos son análisis explícitos, que no sean peticiones de principio o leyes sin advertencias generosas y cláusulas *ceteris paribus*.

Un breve estudio de tres intentos recientes por proporcionar la ley o el análisis apropiado me ayudará a explicar las dificultades. Primero, considérese la siguiente ley de acción propuesta por Paul Churchland:<sup>12</sup>

*si A quiere  $\phi$  y cree que hacer  $x$  es una manera de producir  $\phi$  y que no hay mejor manera de producir  $\phi$ , y A no tiene ningún deseo dominante, sabe cómo  $x$  y es capaz de  $x$ , entonces, A hace  $x$ . [He simplificado la formulación de Churchland.]*

Es claro que si fuese aceptable esta propuesta, podríamos definir ser capaz de  $x$  como el estado de un agente que, cuando son satisfechas

12. «The Logical Character of Action Explanations», pp. 221–222.

las otras condiciones, hace  $x$ . Pero la «ley» es defectuosa, pues en primer lugar, la noción de «deseo dominante» es traicionera, ya que no es claro cómo mostrar que un deseo es dominante excepto por el hecho de que domine. Sin embargo, la dificultad más general es ésta: todas las condiciones pueden ser satisfechas y, sin embargo, no conseguir iniciar una acción. Podría suceder de manera tan simple como ésta: el agente quiere  $\phi$  y cree que hacer  $x$  es la mejor forma de producir  $\phi$  y, sin embargo, él no llega a reunir estas dos cosas; el razonamiento práctico que lo llevaría a concluir que vale la pena hacer  $x$  simplemente puede no darse. No hay más razón para suponer que una persona que tiene razones para actuar, actuará siempre conforme a ellas, que para suponer que una persona que tiene creencias que implican formalmente cierta conclusión, extraerá esa conclusión.

Alvin Goldman ilustra de la siguiente manera el entramado de las tres empresas que he estado estudiando. Primero, define un tipo de acto básico como una especie de acto que realiza un agente si ninguna fuerza externa se lo impide y quiere realizar un acto de ese tipo. El autor prosigue:<sup>13</sup>

si  $x$  es un tipo de acto básico para  $A$  en  $t$  y si  $A$  está en condiciones estándares con respecto a  $x$  en  $t$ , entonces es verdadero el siguiente condicional causal: «Si  $A$  quisiese hacer  $x$  en  $t$ , entonces  $A$  haría  $x$  en  $t$ ». Este enunciado condicional causal [...] implica formalmente el enunciado de que  $A$  es «capaz» de hacer  $x$  en  $t$ . [He alterado el simbolismo de letras.]

Puesto que Goldman hace de querer hacer  $x$  la causa actitudinal de hacer  $x$ , parece que su análisis evita nuestra dificultad. Pero de hecho no lo hace, como lo vemos cuando distinguimos claramente actos tipo y acciones particulares. Querer hacer  $x$  es, para Goldman, un deseo por un tipo de acción, no por una acción particular. Esto está bien; los objetos del querer y del desear son de carácter proposicional y, así, no pueden ser acciones particulares. Pero entonces, es un error decir que querer hacer  $x$  (esto es, realizar una acción de tipo  $x$ ) causa  $x$ , puesto que  $x$ , al no ser un suceso sino un tipo, no puede causarse. Lo que puede causarse es sólo algún suceso que pertenece

13. *A Theory of Human Action*, p. 199.

al tipo. ¿Cuál de éstos ha de causarse? Goldman enuncia una ley que, reformulada, sería como esto: si a una persona no la limitan fuerzas externas, entonces si quiere hacer algo de tipo  $x$ , ella hace algo de tipo  $x$ . Se supone que esto se aplica sólo a actos «básicos» pero, de hecho, no creo que se aplique a ninguna clase interesante de acciones. El problema es que querer hacer algo de tipo  $x$ , aun si el hacer  $x$  está completamente en poder del agente y éste lo sabe, puede no causar que él haga algo de tipo  $x$ . Y un problema mayor es que querer hacer algo de tipo  $x$  puede causar que alguien haga algo de tipo  $x$  y, sin embargo, la cadena causal puede operar de tal manera que el acto no sea intencional.

David Armstrong, en un artículo llamado «Acting and Trying»,<sup>14</sup> es quien más se acerca a ver la naturaleza de la dificultad. Armstrong formula la pregunta de qué se debe añadir a « $A$  intentó hacer  $x$ » a fin de tener condiciones suficientes y necesarias para « $A$  hizo  $x$  intencionalmente». La mayoría de nosotros diría —yo ciertamente lo diría— que el intentar mismo no es necesario en muchos casos, pero esta propuesta, aun cuando central para el interés de Armstrong, en gran medida no es pertinente para el tema actual. Conforme a Armstrong, el que  $A$  haga intencionalmente  $x$  implica formalmente que el intento de  $A$  por hacer  $x$  causó que  $x$  sucediera, pero el intento de  $A$  por hacer  $x$ , aun cuando cause que  $x$  suceda, no prueba que  $A$  hizo  $x$  intencionalmente. La dificultad es que el intento puede producir el efecto deseado de una manera inesperada o no deseada. He aquí un ejemplo de Daniel Bennett. Un hombre puede intentar matar a alguien disparándole. Supongamos que el asesino falla su disparo por una milla, pero el tiro provoca una estampida de una piara de cerdos salvajes que pisotean y matan a la víctima señalada. ¿Queremos decir que el hombre mató a su víctima *intencionalmente*? La cuestión es que no cualquier conexión causal entre actitudes racionalizantes y un efecto deseado bastan para garantizar que producir el efecto deseado fue intencional. La cadena causal debe seguir el tipo de ruta correcto.

Armstrong intenta llenar este hueco diciendo que el efecto deseado debe ser producido por una cadena causal que responda, al

14. Armstrong ha replicado a mis comentarios en una edición posterior de *Philosophical Papers*, mayo de 1975.

menos aproximadamente, al esquema del razonamiento práctico. En el ejemplo de Bennett debemos suponer que el agente tuvo la intención de matar a la víctima jalando el gatillo porque razonó que jalando el gatillo causaría que la pistola disparara, lo que causaría que volara la bala, lo que causaría que la bala penetrara en el cuerpo de la víctima, causando así su muerte. Pero la trama de los sucesos, reflejada por el esquema del razonamiento práctico, no se produjo por la acción de jalar el gatillo. Esto pone en duda la cuestión acerca de si el agente mató intencionalmente a su víctima.

No estoy seguro sobre si esta dificultad pueda superarse o no, pero hay un problema relacionado que Armstrong no considera y que a mí me parece insuperable. Éste es el problema, no de una extraña cadena causal *externa*, sino de cadenas causales *internas*, no estándares o locas. Armstrong, al intentar poner remedio al problema acerca de relaciones causales externas no deseadas, se vio llevado a hablar de la ruta del razonamiento práctico, la manera como interactúan las creencias y los deseos para producir la acción. Y aquí vemos que el análisis de Armstrong, como el que propuse unas cuantas páginas atrás, debe enfrentarse a la cuestión de *cómo* las creencias y los deseos causan acciones intencionales. Las creencias y los deseos que racionalizarían una acción si la causasen de la manera *correcta* —por una ruta de razonamiento práctico, como podríamos intentar decir— pueden causarla de otras maneras. Si es así, la acción no se realizó con la intención que podríamos haber discernido de las actitudes que la causaron. Lo que me desespero por aclarar es la forma como las actitudes deben causar las acciones si han de racionalizar la acción.

Sirva un ejemplo de muestra. Un alpinista podría querer quitarse el peso y peligro de sostener a otro hombre de una cuerda y podría saber que soltar la cuerda podría quitarle el peso y el peligro. Esta creencia y este deseo podrían desconcertarlo tanto como para causar que la soltase y, sin embargo, podría suceder que nunca *eligiese* debilitar su presión, ni hacerlo intencionalmente. Pienso que no ayudará añadir que la creencia y el querer deben combinarse para causar que él quiera soltarla, pues aún quedarán las *dos* preguntas de *cómo* la creencia y el deseo causaron el segundo deseo y *cómo* querer soltarla causó que la soltase.

Algunas páginas atrás intenté analizar «A es libre de hacer  $x$  (o puede hacer  $x$ )» en términos del condicional «Él haría  $x$  intencionalmente si tuviese actitudes que racionalizasen su hacer  $x$ ». Aun si leemos este condicional subjuntivo como si implicara una relación causal, ahora podemos ver que no es adecuado. Si el agente hace  $x$  intencionalmente, entonces su hacer  $x$  lo causan sus actitudes que racionalizan  $x$ . Pero, puesto que puede haber cadenas causales aviesas, no podemos decir que si las actitudes que racionalizarían  $x$  causan que un agente haga  $x$ , entonces hace  $x$  intencionalmente.

Es en gran medida porque no podemos ver cómo completar el enunciado de las condiciones causales de la acción intencional que no podemos decir si, en caso de tenerlo bien, el resultado sería un caso de análisis o una ley empírica para predecir la conducta. Si comenzamos con la sugerencia de que si un hombre tiene creencias y deseos que racionalicen una acción de tipo  $x$ , entonces intencionalmente realiza una acción de tipo  $x$ , es fácil ver que mejoramos las cosas si añadimos al antecedente «y esas creencias y deseos, o el llegar a tenerlos, causan que el agente realice una acción de tipo  $x$ ». Mejorar esta formulación, a su vez, de una manera que eliminase cadenas causales erróneas, eliminaría la necesidad de depender de una apelación abierta a cadenas causales. Simplemente diríamos, dadas estas condiciones (especificadas), que siempre hay una acción intencional de un tipo especificado. Esto, claro está, se entendería como una ley causal, pero no necesitaría mencionar la causalidad. La mención inevitable de la causalidad es un disfraz para la ignorancia; debemos apelar a la noción de causa cuando carecemos de leyes detalladas y precisas. En el análisis de la acción, la mención de la causalidad reduce un poco la distancia que separa el análisis de la ciencia.

Si se produjese la formulación acabada de la ley propuesta, podríamos intentar juzgar su carácter de manera más precisa. Si los términos de las condiciones antecedentes hubiesen de permanecer mentalistas (por ejemplo, por usar conceptos tales como los de creencia y deseo), la ley continuaría pareciendo constitutiva o analítica. Pero el intento de ser más precisos, de eliminar la apelación a nociones que señalan una petición de principio, como deseos «dominantes», pienso que sugeriría que una ley sería tendría que

enunciar las condiciones antecedentes en términos físicos o al menos conductistas. Entonces la ley sería más claramente empírica. Pero el intento por encontrar tal ley concluiría también en un fracaso, aun cuando por razones diferentes.<sup>15</sup>

Debemos considerar un fracaso nuestra búsqueda de un análisis causal de «A es libre de hacer x». ¿Muestra esto que, después de todo, Austin tenía razón? Bien, él tenía razón al rechazar el análisis de Moore (aun cuando, como lo he argumentado, muchas de sus razones eran defectuosas). Y creo que tenía razón al sostener que ningún otro condicional causal podría analizar «A es libre de hacer x». Pero Austin también creía, por tentativo que esto fuera, que había mostrado que ser libre para realizar una acción no puede, de manera alguna, ser un poder causal y, portanto, que es falsa la teoría causal de la acción (él la llama «determinismo», pero no quiero adentrarme en esto). Creo que estaba equivocado en este último aspecto importante. Pues aun cuando no podemos esperar definir o analizar la libertad para actuar en términos de conceptos que identifiquen plenamente las condiciones causales de la acción intencional, no hay ningún obstáculo para la tesis de que la libertad para actuar es un poder causal del agente.

15. En los ensayos 11 y 12 argumento por qué son imposibles las leyes psicofísicas estrictas.



## 5. TENER LA INTENCIÓN

Alguien puede tener la intención de construir una casa de ardillas sin haberse decidido a hacerlo, haber deliberado acerca de ello, haberse formado la intención de hacerlo o haber razonado acerca de ello. Y, a pesar de su intención, podría nunca construir una casa de ardillas, tratar de construir una casa de ardillas o hacer cualquier cosa con la intención de hacer que se construya una casa de ardillas. El mero tener la intención de este tipo, tener la intención que puede ocurrir sin razonamiento práctico, acción o consecuencia, presenta un problema si queremos dar una explicación del concepto de intención que no invoque episodios o actitudes no analizados como querer, actos misteriosos de la voluntad o tipos de causación ajenos a la ciencia.

Cuando se añade la acción a la intención, por ejemplo, cuando alguien une con clavos dos tablas con la intención de construir una casa de ardillas, entonces al principio puede parecer que no surge necesariamente el mismo problema. Somos capaces de explicar qué es lo que sucede en tal caso sin suponer o postular ningún suceso, episodio, actitud o actos extraños o especiales. He aquí cómo lo podemos explicar. Alguien que actúa con cierta intención actúa por una razón; tiene en la mente algo que quiere promover o realizar. Un hombre que une con clavos dos tablas, con la intención de construir una casa de ardillas, debe querer construir una casa de ardillas, o pensar que debe hacerlo (sin duda por otras razones) y debe creer que unir con clavos dos tablas favorecerá su proyecto. Una referencia a otras actitudes, además de querer o pensar que debe, puede ayudar a especificar las razones del agente, pero pa-

rece que debe estar comprendida una actitud positiva o favorable. Cuando hablamos de esta manera de razones, no requerimos que las razones sean buenas. Aprendemos algo acerca de las razones de alguien para iniciar una guerra cuando nos enteramos de que lo hizo con la intención de concluir todas las guerras, aun si sabemos que era falsa su creencia de que iniciar una guerra concluiría todas las guerras. De manera similar, un deseo de humillar a un conocido puede ser la razón que tenga alguien para despreciarlo en una fiesta, aun cuando un observador, en una actitud más normativa, podría pensar que ésa no era razón. La falsedad de una creencia o lo patentemente equivocado de un valor o de un deseo no incapacitan a la creencia o al deseo para proporcionar una razón explicativa. Por otra parte, las creencias y los deseos nos manifiestan las razones del agente para actuar sólo si esas actitudes están relacionadas de manera adecuada a la acción tal como la ve el actor. Para que sirvan como razones de una acción, no es necesario que las creencias y los deseos sean razonables; pero de cualquier modo se introduce un elemento normativo, puesto que la acción debe ser razonable a la luz de las creencias y los deseos (naturalmente, puede no ser razonable a la luz de otras consideraciones).

¿Qué es lo que significa decir que una acción, tal como la ve el agente, es razonable a la luz de sus creencias y sus deseos? Supongamos que un hombre aborda un avión marcado «Londres» con la intención de abordar un avión que se dirija a Londres, Inglaterra. Sus razones para abordar el avión marcado «Londres» las dan su deseo de abordar un avión que se dirija a Londres, Inglaterra, y su creencia de que el avión marcado «Londres» se dirige hacia Londres, Inglaterra. Sus razones explican por qué abordó intencionalmente el avión marcado «Londres». De hecho, el avión marcado «Londres» se dirigía a Londres, Ontario, no Londres, Inglaterra y, así, sus razones no pueden explicar por qué abordó un avión que no se dirigía a Londres, Inglaterra. Pueden explicar por qué abordó un avión que se dirigía a Londres, Ontario, pero sólo cuando se conjugan las razones con el hecho de que el avión marcado «Londres» se dirigía a Londres, Ontario y, claro está, sus razones no pueden explicar por qué abordó intencionalmente un avión que se dirigía

a Londres, Ontario, puesto que no tenía ninguna intención semejante.<sup>1</sup>

La relación entre razones e intenciones puede apreciarse comparando estos enunciados:

- (1) Su razón para abordar el avión marcado «Londres» era que quería abordar un avión que se dirigiera a Londres, Inglaterra y él creía que el avión marcado «Londres» se dirigía a Londres, Inglaterra.
- (2) Su intención al abordar el avión marcado «Londres» era la de abordar un avión que se dirigiera a Londres, Inglaterra.

La primera de estas oraciones implica lógicamente la segunda, pero no a la inversa. La falla de la segunda implicación se debe a dos diferencias entre (1) y (2). Primero, de (2) no es posible reconstruir la actitud favorable específica que se menciona en (1). Dado (2), debe haber alguna actitud favorable adecuada, pero no tiene que ser la de querer. Y, segundo, la descripción de la acción («abordar el avión marcado "Londres"») ocupa un contexto opaco en (1), pero uno transparente en (2). Así, «abordar el avión que se dirige a Londres, Ontario», describe la misma acción que «abordar el avión marcado "Londres"», puesto que el avión marcado «Londres» *era* el avión que se dirigía a Londres, Ontario. Pero la sustitución de «abordar el avión marcado "Londres"» con «abordar el avión que se dirigía a Londres, Ontario» hará falso (1) y dejará verdadero (2). Claro está que la descripción de la intención en (2), como la descripción de los contenidos de la creencia y de la actitud favorable en (1), ocupa un contexto opaco.

Finalmente, hay esta relación entre enunciados con las formas de (1) y de (2): aun cuando (2) no implica formalmente (1), si (2) es verdadero, *algún* enunciado con la forma de (1) es verdadero (con, quizás, otra descripción de la acción y con una actitud favorable y una creencia apropiadas en los lugares adecuados). El enunciado (1), a diferencia del (2), debe describir la acción del agente de manera

1. Considero que el «intencionalmente» gobierna toda la frase «abordó un avión que se dirigía a Londres, Ontario». Conforme a una lectura alternativa, sólo el abordaje sería intencional. De manera similar, en el (1) que sigue, sus razones se extienden al rótulo en el avión.

que haga claro un sentido en el que la acción era razonable a la luz de las razones del agente. Así, podemos decir, si un agente hace *A* con la intención de hacer *B*, que hay alguna descripción de *A* que revela que la acción es razonable a la luz de las razones que el agente tenía para realizarla.

¿Cuándo es razonable una acción (descrita de una manera particular) a la luz de las creencias y de las actitudes favorables específicas? Una manera de abordar el asunto es mediante una versión bastante abstracta del razonamiento práctico. No podemos suponer que siempre que un agente actúa intencionalmente recorre un proceso de deliberación o de razonamiento, ordena evidencia y principios y extrae conclusiones. Sin embargo, si alguien actúa con una intención, debe tener actitudes y creencias de las que, si hubiese estado consciente de ellas y hubiese tenido tiempo, *podría* haber razonado que su acción era deseable (o tenía algún otro atributo positivo). Si podemos caracterizar el razonamiento que serviría, habríamos descrito, en efecto, las relaciones lógicas entre descripciones de creencias y deseos y la descripción de una acción, cuando la primera da las razones con las que se realizó la última. Entonces hemos de imaginar que las creencias y los deseos del agente le proporcionan las premisas de un argumento. En el caso de la creencia es claro, de inmediato, cuál es la premisa. Tomemos un ejemplo: alguien le añade salvia al guisado con la intención de mejorar su sabor. Así, su premisa correspondiente es: añadirle salvia al guisado mejorará su sabor.

La actitud favorable del agente es, quizás, un deseo o un querer; supongamos que quiere mejorar el sabor del guisado. Pero, ¿cuál es la premisa correspondiente? Si buscásemos la proposición hacia la que se dirige su deseo, la proposición que quiere verdadera, sería algo como: él hace algo que mejora el sabor del guisado (de manera más breve: él mejora el sabor del guisado). Sin embargo, ésta no puede ser su premisa, pues nada interesante se sigue de las dos premisas: añadir salvia al guisado mejorará su sabor y el agente mejora el sabor del guisado. El problema es que se ha dejado fuera la actitud de *aprobación* que tiene el agente hacia la segunda proposición. No puede volverse a introducir haciendo que la premisa sea «el agente quiere mejorar el sabor del guisado»; no queremos una *descripción* de su deseo, sino una *expresión* del mismo de una forma en la que

él podría usarlo para llegar a una acción. La expresión natural de su deseo es, según me parece, de forma evaluativa; por ejemplo, «es deseable mejorar el sabor del guisado» o «debo mejorar el sabor del guisado». Podemos suponer que diferentes actitudes favorables se expresan con otras palabras evaluativas en lugar de «deseable».

No hay ninguna prueba breve de que las oraciones evaluativas expresen deseos y otras actitudes favorables a la manera como la oración «la nieve es blanca» expresa la creencia de que la nieve es blanca. Pero la siguiente consideración nos ayudará, quizás, a mostrar lo que está en juego. Si alguien que sabe español dice, honestamente, «la nieve es blanca», entonces él cree que la nieve es blanca. Si mi tesis es correcta, alguien que dice honestamente «es deseable que yo deje de fumar» tiene una actitud favorable hacia su dejar de fumar. Siente alguna inclinación hacia ello; de hecho, lo hará si nada se interpone, sabe cómo hacerlo y no tiene ningún valor o deseo en contra. Dado este supuesto, es razonable generalizar: si juicios explícitos de valor representan actitudes favorables, todas las actitudes favorables pueden expresarse mediante juicios de valor que sean al menos implícitos.

Esta última estipulación nos permite dar una versión uniforme de actuar con una intención. Si alguien realiza una acción de tipo *A* con la intención de realizar una acción de tipo *B*, entonces debe tener una actitud favorable hacia acciones de tipo *B* (lo que puede expresarse de la forma: una acción de tipo *B* es buena (o tiene algún otro atributo positivo)) y una creencia de que al realizar una acción de tipo *A* estará (o probablemente estará) realizando una acción de tipo *B* (la creencia puede expresarse de la manera obvia). La expresión de la creencia y del deseo implican formalmente que las acciones de tipo *A* son o probablemente serán buenas (o deseables, justas, debidas, etc.). La descripción de la acción que proporciona la frase sustituida por «*A*» da la descripción en la cual el deseo y la creencia racionalizan la acción. Así que para devolver las cosas a nuestro ejemplo, el deseo de mejorar el sabor del guisado y la creencia de que añadir salvia al guisado mejorará su sabor sirven para racionalizar una acción descrita como «añadir salvia al guisado» (esta versión más o menos estándar del razonamiento práctico se modificará radicalmente en un momento).

Debe haber tales creencias y deseos racionalizantes si una acción ha de hacerse por una razón pero, claro está, la presencia de tales creencias y deseos, cuando la acción se hace, no basta para asegurar que lo que se hizo se hizo con la intención apropiada o siquiera con intención alguna. Alguien podría querer un guisado sabroso, creer que la salvia lo lograría y ponerle salvia pensando que era perejil; o ponerle salvia porque le empujaron la mano. Así que debemos añadir que el agente puso la salvia por sus razones. Este «por» es una fuente de problemas; según creo, y esto lo he argumentado ampliamente, implica la noción de causa. Pero no servirá cualquier relación causal, puesto que un agente podría tener actitudes y creencias que racionalizarían una acción y podrían ser la causa de que la realizara y, sin embargo, debido a alguna anomalía en la cadena causal, la acción no sería intencional en el sentido esperado o, quizás, en ningún sentido.<sup>2</sup>

Concluimos, entonces, con esta versión incompleta e insatisfactoria de actuar con una intención: una acción se realiza con cierta intención si la causan, de la manera correcta, las actitudes y las creencias que la racionalizan.<sup>3</sup>

Si esta versión es correcta, entonces actuar con una intención no requiere que haya algún misterioso acto de la voluntad, alguna actitud especial o un episodio de volición; pues la versión sólo necesita deseos (u otras actitudes favorables), creencias y las acciones mismas. Ciertamente está la relación, aún por analizar, causal o de otra naturaleza, entre éstas, pero no es una entidad embarazosa que tenga que añadirse al bagaje del mundo. Es verdad que no habríamos mostrado cómo *definir* el concepto de actuar con una intención; la reducción no es definicional sino ontológica. Pero la reducción ontológica, si tiene éxito, basta para responder a muchas perplejidades acerca de la relación entre la mente y el cuerpo y para explicar la posibilidad de la acción autónoma en un mundo de causalidad.

Esto me devuelve al problema que mencioné al principio, pues la estrategia que parece funcionar para actuar con una intención no tiene ninguna aplicación obvia al mero pretender, esto es, a tener una

2. Véase el ensayo 4.

3. Hasta aquí dejó las cosas el ensayo 1. Cuando lo escribí creí que sería posible caracterizar en términos no circulares «la manera correcta».

intención no necesariamente acompañada por la acción. Si alguien cava una trampa con la intención de atrapar un tigre, es quizás plausible que ninguna entidad, acto, suceso o disposición en absoluto corresponda a la frase nominal «la intención de atrapar un tigre» —esto es lo que nuestra revisión nos ha llevado a esperar. Pero no es probable que si un hombre tiene la intención de atrapar un tigre, su intención no sea un estado, disposición o actitud de algún tipo. Sin embargo, si esto es así, es absolutamente increíble que este estado o actitud (y el suceso o acto relacionado de *formar una intención*) no vaya a desempeñar ningún papel en el actuar con una intención. Nuestra falta de habilidad para dar una versión satisfactoria de meramente tener la intención, sobre la base de nuestra versión de la acción intencional, se refleja así sobre la versión misma de acción intencional. Y creo que se verá que la versión que he delineado es incompleta cuando tengamos un análisis adecuado de meramente tener la intención.

Claro está que realizamos muchas acciones intencionales sin formar una intención de realizarlas y, con frecuencia, la acción intencional no es precedida por una intención. Así, no sería sorprendente si algo estuviese presente en meramente tener la intención que no esté siempre presente en la acción intencional. Pero sería sorprendente si ese elemento extra fuese extraño a nuestra comprensión de la acción intencional. Pues considérese alguna acción simple, como escribir la palabra «acción». Algunos segmentos temporales de esta acción son también acciones; por ejemplo, primero escribo la letra «a». Esto lo hago con la intención de iniciar una acción que no se completará hasta que haya escrito el resto de la palabra. Es difícil ver cómo la actitud hacia el acto completo que tengo mientras escribo la letra «a» difiere de la mera intención que haya podido tener un momento antes. Ciertamente mi intención se ha comenzado a realizar ahora, pero ¿por qué eso necesariamente ha de cambiar mi actitud? Parece que en cualquier acción intencional que tome mucho tiempo o comprenda pasos preparatorios, debe estar presente algo como meramente tener la intención.

Comenzamos con meramente tener la intención —tener la intención sin deliberación consciente o una consecuencia externa— porque no había espacio para dudar de que tener la intención es un estado o suceso separado de la acción que se pretende o de las razo-

nes que mueven a la acción. Una vez que se reconoce la existencia de meramente tener la intención, no hay razón para no conceder que está presente la intención exactamente del mismo tipo cuando sucede la acción que se pretendía. Así que, en lo que sigue, aun cuando parezca que me concentro en el caso muy especial de intenciones no satisfechas, el tema, de hecho, es cualquier caso de tener la intención, esto es, el tener la intención abstraído de un contexto que pueda incluir cualquier grado de deliberación y cualquier grado de éxito en la ejecución. Meramente tener la intención muestra tan sólo que ahí hay algo por abstraer.

El éxito que hayamos tenido al enfrentarnos al concepto de acción intencional provino de tratar el discurso de la intención con la que se hace una acción como discurso de creencias, deseos y acción. Esto sugiere que intentemos tratar las meras intenciones —el tener intenciones abstraído de los resultados normales— como acciones, creencias o actitudes favorables de algún tipo. El resto de este ensayo considera estas posibilidades.

*¿Meramente tener la intención es una acción?* Podría objetarse que tener la intención de hacer algo no es un cambio o un suceso de algún tipo y, así, no puede ser algo que haga un agente. Pero esta objeción se refuta ajustando la tesis; diríamos que la acción es formar una intención en tanto que meramente tener la intención es el estado de un agente que ha formado una intención (y no ha cambiado de opinión). Así, todo el peso se pone en la idea de formar una intención. Se dirá que la mayoría de las intenciones no se forman, al menos si formar una intención requiere una deliberación o una decisión consciente. Entonces lo que necesitamos es el concepto más amplio y más neutral de llegar a tener una intención —un cambio que puede darse de manera tan lenta o inadvertida que el agente no pueda decir cuándo sucedió. De cualquier forma, es un suceso y podemos decidir llamarlo una acción o, al menos, algo que hace el agente.

No veo ninguna razón para rechazar esta propuesta; lo peor que puede decirse de ella es que es muy poco ilustrativa. El estado de intención es sólo lo que resulta de llegar a tener una intención —pero, ¿qué tipo de estado es? El llegar a tener una intención podríamos conectarlo con deseos y creencias, como lo hicimos con otras acciones intencionales (nuevamente con una cadena causal que funcione



«de manera correcta»). Pero la historia no tiene la calidad sustancial de la versión de la acción intencional, porque la acción pretendida no es familiar u observable, ni siquiera para el mismo agente.

Otro enfoque se centra en actos de habla públicos. *Decir* que uno tiene la intención de hacer algo o que uno lo hará es, de manera innegable, una acción y tiene algunas de las características de formar una intención. Decir, en las circunstancias apropiadas, que uno tiene la intención de hacer algo o que uno lo hará, puede comprometerlo a uno a hacerlo; si el hecho no se sigue, es apropiado pedir una explicación. De hecho, identificar el decir que uno tiene la intención de hacer algo con formar una intención sería apoyar una especie de teoría realizativa de la intención; así como decir que uno promete hacer algo puede ser prometer hacerlo, decir que uno tiene la intención de hacer algo puede ser pretender (o formar la intención de) hacerlo. Claro está que uno puede formar una intención sin decir nada en voz alta, pero puede llenar este hueco la noción de hablarse a sí mismo, «decirlo para sus adentros». <sup>4</sup> Una variante teórica sería asemejar (o identificar) formar una intención con dirigirse un mandato a uno mismo.

Pienso que es fácil ver que formar una intención es muy distinto de decir algo, aunque sea a uno mismo, y que tener la intención de hacer algo es muy distinto de haber dicho algo. Pues, por un lado, el carácter realizativo de los mandatos y de las promesas que hacen que ciertos actos de habla sean sorprendentemente importantes depende de convenciones muy específicas y no hay tales convenciones gobernando la formación de intenciones. El prometer comprende asumir una obligación; pero incluso si hay obligaciones para con uno mismo, el tener la intención normalmente no crea ninguna. Si un agente no hace lo que tenía la intención de hacer, normalmente no se debe a sí mismo ninguna explicación o disculpa, especialmente si simplemente cambió de opinión; sin embargo, éste es justamente el caso que pide una explicación o una disculpa cuando se ha quebrantado una promesa hecha a otro. Un mandato puede desobedecerse, pero sólo cuando está vigente. Pero si un agente no hace lo que tenía la intención de hacer porque ha cambiado de opinión, la intención

4. Véase *Mental Acts* de P.T. Geach.

original ya no está vigente. Quizás baste, para desacreditar estas teorías, señalar que el prometer y el mandar, tal como usualmente los entendemos, son, necesariamente, realizaciones públicas, en tanto que formar una intención no lo es. Formar una intención puede ser una acción, pero no es una realización, y tener una intención no es, en general, el resultado de una.

Nada de esto es negar que decir «tengo la intención de hacerlo» o «lo haré» es muy similar o, en ocasiones, idéntico a prometer hacerlo. Si digo cualquiera de estas cosas en el contexto correcto, le doy derecho a un oyente a creer que yo creo que lo haré. Quizás una forma más simple de formularlo es ésta: si digo en ciertas condiciones «tengo la intención de hacerlo», «lo haré» o «prometo hacerlo», entonces *me represento a mí mismo* como si creyera que lo haré. Puedo no creer que lo haré, puedo no tener la intención de que mi escucha crea que lo haré, pero le estoy dando un fundamento para que se queje si no lo hago. Estos hechos sugieren que si no sólo digo «tengo la intención» o «lo haré» de tal manera que me represente a mí mismo como si creyera que lo haré, sino que también soy sincero, entonces mi sinceridad garantiza tanto que tengo la intención de hacerlo como que creo que lo haré. Una línea de argumento semejante a ésta ha llevado a muchos filósofos a sostener que tener la intención de hacer algo implica formalmente creer que uno lo hará y ha llevado a unos cuantos filósofos a la opinión más extrema de que tener la intención de hacer algo es idéntico a un tipo de creencia de que uno lo hará.

*¿Tener la intención de actuar es una creencia de que uno lo hará?*

El argumento que acabo de esbozar ni siquiera muestra que tener la intención implique creencia. El argumento prueba que un hombre que, en ciertas condiciones, sinceramente dice «tengo la intención de hacerlo» o «lo haré» debe creer que lo hará. Pero puede ser el decirlo y no la intención lo que implique la creencia. Y pienso que podemos ver que esto es así. El problema es que hemos pedido que la noción de sinceridad ejecute dos tareas. Comenzamos considerando casos en los que al decir «tengo la intención de» o «lo haré», le doy derecho a un escucha a creer que lo hará. Y aquí es obvio que si soy sincero, si las cosas son como las represento, entonces debo creer que lo haré. Pero es un supuesto no apoyado por el argumento de

que en cualquier momento en el que sinceramente diga que tengo la intención de hacer algo, debo creer que lo haré, pues la sinceridad, en este caso, tan sólo requiere que yo sepa que tengo la intención de hacerlo. Estamos de acuerdo en que hay casos en los que la sinceridad en el emisor de «tengo la intención de» requiere que quien lo emite crea que lo hará, pero el argumento requiere que estos casos incluyan todos aquellos en los que el hablante sabe o cree que tiene la intención de hacerlo.

Una vez que hemos distinguido la cuestión de cómo la creencia se introduce en los reconocimientos de intención de la cuestión de cómo la creencia se introduce en la intención, nos debe saltar a la vista lo dudosa que es la última conexión.

Es un error suponer que si un agente está haciendo algo intencionalmente, debe saber que lo está haciendo. Pues supongamos que un hombre está escribiendo su testamento con la intención de velar por el bienestar de sus hijos. Él puede dudar acerca de su éxito y seguir así hasta su muerte; sin embargo, al escribir su testamento, él puede, de hecho, estar velando por el bienestar de sus hijos y, si es así, ciertamente lo está haciendo de manera intencional. Algunos escépticos pueden pensar que este ejemplo falla porque se niegan a conceder que un hombre puede *ahora* velar por el bienestar de sus hijos si ese bienestar incluye sucesos aún por acontecer. Así que he aquí otro ejemplo: al escribir con fuerza sobre este papel puedo tener la intención de producir diez copias al carbón legibles. No sé o no creo con confianza que esté teniendo éxito. Pero, si estoy produciendo diez copias al carbón legibles, ciertamente lo estoy haciendo intencionalmente. Estos ejemplos no prueban que el mero tener la intención no pueda implicar creencia, pues los ejemplos comprenden el actuar con una intención. Sin embargo, es difícil imaginar que la propuesta no se aplique al meramente tener la intención. Mientras escribe su testamento, el hombre no sólo está actuando con la intención de velar por el bienestar de sus hijos, sino que también tiene la intención de velar por el bienestar de sus hijos. Si puede dudar acerca de si ahora hace lo que tiene la intención de hacer, de seguro puede dudar acerca de si hará lo que tiene la intención de hacer.

La tesis de que tener la intención implica creer se defiende en ocasiones sosteniendo que las expresiones de intención son, en general, incompletas o elípticas. Así, el hombre que escribe su testamento

debe describirse como si tuviera la intención de tratar de velar por el bienestar de sus hijos, no como si tuviera la intención de velar por él; y el hombre con el papel carbón tan sólo está teniendo la intención de tratar de producir sus copias. Las frases suenan erróneas; nosotros estaríamos mucho más dispuestos a decir que *está* tratando e intenta hacerlo. Pero cuando la acción está completamente en el futuro, en ocasiones reconocemos que tenemos la intención de tratar y vemos esto como más preciso que el mero enunciado de intención cuando el resultado es suficientemente dudoso. Sin embargo, no creo que el cargo de elipsis pueda usarse para defender la tesis general.

Sin duda, muchas intenciones tienen forma condicional —se tiene la intención de hacer algo sólo si están satisfechas ciertas condiciones— y, sin duda, por una u otra razón, con frecuencia suprimimos la mención de las condiciones. Así que son comunes los enunciados elípticos de tener la intención. Grice nos da el siguiente diálogo:

X. Tengo la intención de ir a ese concierto el martes.

Y. Lo gozarás.

X. Podría no estar ahí.

Y. Me temo que no te entiendo.

X. La policía me va a formular unas preguntas molestas el martes por la tarde y puedo estar en la cárcel el martes por la noche.

Y. Entonces, para comenzar, deberías haber dicho «tengo la intención de ir al concierto si no estoy en cárcel» o, si querías ser más reservado, algo como esto: «probablemente iré», «espero ir», «tengo el propósito de ir» o «tengo la intención de ir si puedo».<sup>5</sup>

Grice no habla aquí de elipsis, pero piensa que este ejemplo y otros similares conforman un caso fuerte a favor de la tesis de que «X tiene la intención de hacer A» es verdadera cuando «tiene la intención» se usa en el sentido *estricto*, sólo si X está seguro de que hará A. El hombre del ejemplo debe tratar de hacer *algo* y, así, si supiésemos lo que era, podríamos decir que su observación «tengo la intención de ir al concierto» era una elipsis de lo que habría dicho si hubiese usado «tengo la intención» en el sentido estricto. ¿Qué es lo que habría dicho? «Espero ir» no es

5. H.P. Grice, «Intention and Uncertainty», pp. 4–5.

más preciso acerca de la intención, pues no declara ninguna intención; lo mismo sucede con «pretendo ir» y con «probablemente iré»; «tengo la intención de ir si puedo» es vago y general dada la particularidad de las dudas de *X*, pero parece que hay algún error peor aquí. Pues si un agente no puede tener la intención de lo que cree que es imposible, entonces no asevera ni más ni menos diciendo «tengo la intención de ir si puedo» que lo que haría diciendo «tengo la intención de hacerlo». ¿Qué pasaría con «tengo la intención de ir al concierto si no estoy en la cárcel»? Intuitivamente esto se acerca más a comunicar la verdad acerca de la situación tal como *X* la ve. Pero ¿es *literalmente* más precisa? Es difícil ver cómo. Conforme a la tesis a favor de la cual argumenta Grice, si *X* dijese en sentido estricto y honestamente «tengo la intención de ir al concierto», él daría a entender que creía que estaría ahí. Si *X* dijese en el sentido estricto y honestamente «tengo la intención de ir al concierto si no estoy en la cárcel», él daría a entender que creía que estaría en el concierto si no estuviese en la cárcel. Ahora bien, obviamente la primera creencia implica la segunda pero no es implicada por ésta y, así, una expresión de la segunda creencia hace una aseveración menor y puede pensarse que es más precisa. Claro está que la aseveración más fuerte no puede, por su contenido, conducir a *Y* a error acerca de lo que hará *X*, pues sea que *X* diga que estará en el concierto o sólo que estará ahí si no está en la cárcel, tanto *X* como *Y* saben que *X* no estará en el concierto si está en la cárcel. Donde se podría engañar *Y*, si es verdadera la tesis que estamos examinando, es con respecto a lo que *X* cree que hará y, por tanto, a lo que tiene la intención de hacer. Pues, conforme a la tesis, «tengo la intención de ir al concierto si no estoy en la cárcel» implica una creencia más débil que «tengo la intención de ir al concierto». Si esto es correcto, aún se tendrá mayor precisión con más condiciones, ya que *X* tampoco cree que estará en el concierto si cambia de parecer o si se lo impide algo distinto a la prisión. Así, cada vez más nos vemos llevados hacia el casi vacío «tengo la intención de hacerlo si nada me lo impide, si no cambio de parecer, si nada adverso sucede». Esto no nos dice casi acerca de lo que el agente cree acerca del futuro o sobre lo que de hecho hará.

Creo que X habló correctamente y con precisión, pero engañosamente, cuando dijo «tengo la intención de ir al concierto». Él pudo haber corregido la impresión y aún sostener la precisión diciendo «ahora tengo la intención de ir al concierto, pero ya que me pueden encarcelar, podré no estar ahí». Un hombre que dice «tengo la intención de estar ahí, pero podría no estar ahí» no se contradice a sí mismo; a lo sumo, será inescrutable hasta que diga más. Debemos darnos cuenta de que hay algo errado con la idea de que la mayoría de los enunciados de intención, hasta que nuestras dudas los matizan, son elípticos acerca de lo que de hecho haremos cuando nos demos cuenta de que no hay ningún método *general* satisfactorio para aportar el enunciado más preciso del cual era representante el enunciado original. Y la razón es clara: no puede haber ninguna lista finita de cosas que pensemos que pudieran impedirnos hacer aquello de lo que tenemos la intención o de circunstancias que podrían causar que nos detuviéramos. Si estamos razonablemente seguros de que algo nos impedirá actuar, esto quizás no frustré la intención, pero si simplemente estamos inciertos, como sucede con frecuencia, la intención no necesariamente se mella. Podemos tener claridad acerca de lo que tenemos la intención de hacer aun cuando estemos en la oscuridad acerca de los detalles y, por tanto, de los escollos. Si esto es así, ser más precisos acerca de lo que tenemos la intención de hacer no puede ser una cuestión de ser más precisos acerca de lo que creemos que lograremos.

Hay intenciones condicionales genuinas pero no pienso que se presenten en la forma «trataré de hacerlo si puedo» o «si no cambio de opinión». Las intenciones condicionales genuinas son apropiadas cuando consideramos explícitamente qué hacer en diversas contingencias; por ejemplo, alguien puede tener la intención de irse a casa temprano de una fiesta si la música es demasiado fuerte. Si preguntamos por la diferencia entre condiciones que realmente hacen más preciso el enunciado de una intención y condiciones dudosas como «si puedo», «si nada se presenta» o «si no cambio de opinión», me parece claro que la diferencia es ésta: las condiciones *bona fide* son aquellas que son razones para actuar que son contemporáneas con la intención. A alguien puede no gustarle la música fuerte ahora y esto puede ser por lo que tiene ahora la intención de irse de la fiesta a casa

temprano si la música es muy fuerte. El que no pueda irse temprano a casa no es una razón a favor o en contra de que se vaya a casa temprano y, así, no es una condición pertinente para una intención, aunque si cree que no puede hacerlo, eso puede impedirle tener la intención. Cambiar de opinión es un caso complicado pero, en general, nadie está dispuesto a considerar un posible cambio futuro de intención como una razón para modificar su intención actual, a menos que piense que el cambio futuro lo producirá algo que ahora consideraría como una razón.

El contraste que ha surgido entre las circunstancias que en ocasiones aceptamos que condicionan nuestras intenciones y las circunstancias que aceptaríamos si las intenciones implicasen la creencia de que haremos lo que tenemos la intención de hacer, me parece que indica muy concluyentemente que no necesariamente creemos que haremos aquello que tenemos la intención de hacer y que no enunciamos nuestras intenciones de manera más precisa haciéndolas condicionales para todas las circunstancias en cuya presencia pensamos que actuaríamos.

Estas últimas consideraciones apuntan al argumento más fuerte en contra de la identificación de meramente tener la intención con la creencia de que uno hará aquello de lo que tiene la intención. Éste es que las razones para tener la intención de hacer algo son, en general, muy distintas de las razones para creer que uno lo hará. He aquí por qué tengo la intención de arrizar la mayor: veo que viene una turbonada, quiero impedir que zozobre el bote y creo que arrizar la mayor impedirá que zozobre el bote. Yo formularía mis razones para intentar arrizar la mayor de esta manera: viene una turbonada, sería vergonzoso que el bote zozobrar y arrizar la mayor impedirá que el bote zozobre. Pero, por sí mismas, estas razones para intentar arrizar la mayor no me dan ninguna razón para creer que arrizaré la mayor. Dados supuestos adicionales, por ejemplo, que la proximidad de una turbonada es una razón para creer que yo creo que una turbonada está próxima y que lo vergonzoso de que zozobre el bote puede ser una razón para creer que yo quiero impedir que zozobre el bote y dado que tengo estas creencias y deseos, puede ser razonable suponer que tengo la intención de arrizar la mayor y que de hecho lo haré. Así que puede haber una conexión débil

entre las razones de los dos tipos, pero no son en manera alguna idénticas (las razones individuales pueden ser las mismas, pero un muy pequeño conjunto natural de razones que apoyan la intención de actuar no puede ser un conjunto que apoya la creencia de que el acto se llevará a cabo).

Con frecuencia se sostiene que una intención es una creencia a la que no se ha llegado por razonamiento a partir de la evidencia o que una intención es una creencia de una acción futura propia que de alguna otra manera difiere, en su origen, de una predicción ordinaria. Pero tales aseveraciones no ayudan a la tesis. Cómo llegó alguien a una creencia, qué razones daría para apoyarla, qué mantiene su fe, simplemente son asuntos no pertinentes a la cuestión de qué constituyen razones a favor de la creencia; los primeros sucesos son accidentes que le suceden a una creencia y no pueden cambiar su papel lógico sin hacerla una nueva creencia.

*¿Intentar hacer algo es lo mismo que querer hacerlo?* Es claro que las razones para intentar hacer algo son muy similares a las razones para la acción, y ciertamente a uno se le podría decir que son exactamente las mismas excepto por el tiempo. Como lo dijo John Donne, «querer implica dilación», pero podemos reducir la dilación a un momento. Estoy escribiendo la letra «a» de «acción» y tengo la intención de escribir la letra «c» tan pronto como termine con la «a». La razón por la que tengo la intención de escribir la letra «c» tan pronto como termine con la «a» es que quiero escribir la palabra «acción» y sé que, para hacer esto, debo escribir la letra «c» después de escribir la letra «a». ¡Ahora he terminado la «a» y he comenzado la «c»! ¿Cuál es mi razón para escribir la «c»? Es que quiero escribir la palabra «acción» y sé que para hacer esto debo. . . Hasta aquí las razones suenan idénticas pero, si observamos con mayor detenimiento, surgirá una diferencia. Cuando estoy escribiendo la «a», tengo la intención de escribir la «c» en un momento y parte de mi razón es que creo que este momento se perfila en el futuro inmediato; cuando estoy escribiendo la «c» mis razones incluyen la creencia de que *ahora* es el tiempo de escribir la «c», si he de escribir «acción», como lo quiero. En ocasiones, Aristóteles pasa por alto esta diferencia y como resultado de esto dice cosas que suenan fatuas. Él está dispuesto a dar como un ejemplo de razonamiento práctico



algo como lo siguiente: quiero estar caliente, creo que una casa me mantendrá caliente, de inmediato construyo una casa. Una doctrina importante es que la conclusión de un razonamiento práctico puede ser una acción; también es importante que la conclusión puede ser la formación de una intención de hacer algo en el futuro.

Ahora me gustaría llamar la atención hacia un aspecto que parece producir una dificultad en este cuadro de lo que pueda ser formar una intención. Considérese, nuevamente, un caso de acción intencional. Quiero comer algo dulce, esto es, sostengo que es deseable que yo coma algo dulce. Creo que este caramelo es dulce y, así, el que yo coma este caramelo será un caso de que yo coma algo dulce y concluyo que es deseable que yo coma este caramelo. Puesto que nada se interpone en el camino, como el caramelo —la conclusión es la acción. Pero esto también significa que yo podría expresar la conclusión usando una referencia demostrativa a la acción: «Esta acción mía, este comer yo este caramelo ahora, es deseable». Lo que parece tan importante acerca de la posibilidad de una referencia demostrativa a la acción es que es un caso en el que tiene sentido acoplar directamente un juicio de valor a una acción. Mi razón evaluativa para actuar fue «es deseable que yo coma algo dulce». Pero, claro está, esto no puede significar que cualquier acción mía, la que sea, que sea comer algo dulce, es algo que tiene sentido hacer —mi juicio tan sólo trata con acciones en tanto que sean consumidoras de dulce. Algunas de esas acciones, incluso todas ellas, pueden tener muchos otros aspectos indeseables. Es tan sólo cuando llego a una acción real cuando tiene sentido juzgarla, en su totalidad, como deseable o no; hasta llegar a ese momento no había ningún objeto con el que yo estuviese en contacto para juzgar. Claro está que aún puedo decir de la acción completada que es deseable en tanto que sea esto o aquello, pero al elegir realizarla fui más allá de esto; mi elección representaba, o quizás era, un juicio de que la acción misma era deseable.

Y ahora el problema acerca de meramente tener la intención es que no hay ninguna acción para juzgarla simplemente como buena o deseable. Todo lo que podemos juzgar en la etapa de meramente tener la intención es la deseabilidad de acciones de cierto tipo y las acciones de cierto tipo se juzgan generalmente sobre la base del aspecto que define el tipo. Sin embargo, tales juicios no siempre

conducen a una acción razonable, o estaríamos comiendo todo lo dulce sobre lo que pudiésemos poner nuestras manos (como lo ha señalado Anscombe).<sup>6</sup>

El paso más importante para aclarar estos asuntos es establecer una firme distinción entre el tipo de juicio que corresponde a un deseo, como querer comer algo dulce, y el tipo de juicio que puede ser la conclusión de un razonamiento práctico —que puede corresponder a una acción intencional.<sup>7</sup> Con frecuencia se considera que el primer tipo de juicio tiene la forma de una ley: es deseable cualquier acción que sea el comer algo dulce. Si el razonamiento práctico es deductivo, esto es lo que deberíamos esperar (y parece que es como Aristóteles y Hume, por ejemplo, consideraron el razonamiento práctico). Pero hay una objeción fundamental a esta idea, como puede verse cuando consideramos una acción que a la vez tiene un aspecto deseable y uno indeseable. Pues supongamos que la expresión proposicional de un deseo de comer algo dulce fuese un condicional cuantificado universalmente. Al sostener que es deseable comer algo dulce, podemos también sostener que es indeseable comer algo venenoso. Pero uno y el mismo objeto puede ser dulce y venenoso y, así, una y la misma acción puede ser comer algo dulce y algo venenoso. Nuestros principios evaluativos, que parecen consistentes, pueden, entonces, llevarnos a concluir que la misma acción es a la vez deseable e indeseable. Si las acciones indeseables no son deseables, hemos derivado una contradicción de premisas todas las cuales son plausibles. La cura es reconocer que hemos asignado la forma equivocada a los principios evaluativos. Si son juicios en el sentido de que una acción es buena (o deseable, etc.) *en la medida en que* tiene cierta característica, entonces no

6. G.E.M. Anscombe, *Intention*, p. 59. [Hay traducción al español: *Intención*, trad. Ana Isabel Stellino, Instituto de Investigaciones Filosóficas-UNAM/Paidós, México/Barcelona, 1991.]

7. Ningún peso ha de dársele a la palabra «juicio». Aquí estoy considerando la *forma* de las proposiciones que expresan deseos y otras actitudes. No supongo que alguien que quiera comer algo dulce necesariamente *juzgue* que sería bueno comer algo dulce; quizás podamos decir que él *sostiene* que su comer algo dulce tiene alguna característica positiva. Al distinguir entre las expresiones proposicionales de actitudes espero señalar diferencias entre las actitudes.

deben interpretarse de tal manera que funcione la separación<sup>8</sup> o nos encontraremos concluyendo que la acción es simplemente deseable cuando todo lo que está garantizado sea la conclusión de que es deseable en cierto respecto. Llamemos juicios *prima facie* a juicios de que las acciones son deseables en tanto tengan cierto atributo.

Los juicios *prima facie* no pueden asociarse directamente con las acciones, pues no es razonable realizar una acción tan sólo porque tenga una característica deseable. Es una razón para actuar el que se crea que la acción tiene alguna característica deseable, pero el hecho de que se realiza la acción representa un nuevo juicio de que la característica deseable era suficiente para actuar con base en ella —que no pesaban más otras consideraciones. Por tanto, el juicio que corresponde a la acción, o es, quizás, idéntico a ella, no puede ser un juicio *prima facie*; debe ser un juicio global o incondicionado que, si lo expresáramos en palabras, tendría una forma como «esta acción es deseable».

Ahora puede verse que nuestro tratamiento anterior del actuar con una intención era engañoso o al menos incompleto en un aspecto importante. Las razones que determinan la descripción en la cual se tiene la intención de actuar no nos permiten *deducir* simplemente que vale la pena realizar la acción; todo lo que podemos deducir es que la acción tiene un rasgo que habla a su favor. Sin embargo, esto basta para permitirnos dar la intención con la cual se realizó la acción. Lo que es engañoso es que las razones que entran en esta explicación no constituyen, generalmente, todas las razones que consideró el agente al actuar y, así, conocer la intención con la que alguien actuó no nos permite reconstruir su razonamiento real. Pues podemos no saber cómo pasó el agente de sus deseos y otras actitudes —sus razones *prima facie*— a la conclusión de que cierta acción era deseable.<sup>9</sup>

En el caso de la acción intencional, al menos cuando la acción es de duración breve, nada parece impedir una identificación aritotélica de la acción con un juicio de cierto tipo —un juicio global, incondicional, de que la acción es deseable (o tiene alguna otra carac-

8. La regla de *modus ponens*. [N. del T.]

9. En el ensayo 2 he dicho más acerca de la forma de los juicios evaluativos *prima facie* y de la importancia de distinguirlos de los juicios incondicionales.

terística positiva). La identificación de la acción con la conclusión de un razonamiento práctico no le es esencial a la tesis que apoyo, pero el hecho de que pueda hacerse explica por qué, en nuestra versión original de la acción intencional, permanecía oculto lo que se necesitaba para relacionarla con la pura intención.

Ahora sugiero que, en el caso de la pura intención, la intención es simplemente un juicio global. Formar una intención, decidir, elegir y deliberar son diversos modos de llegar al juicio, pero es posible llegar a tener tal juicio o tal actitud sin que se aplique ninguno de estos modos.

Permítaseme desarrollar esta sugerencia y tratar de defenderla de algunas objeciones. Unas cuantas páginas antes señalé que un juicio global tiene sentido sólo cuando hay una acción presente (o pasada) que se conoce de manera directa. De otra manera (argüí) el juicio debe ser general, esto es, debe cubrir todas las acciones de cierto tipo y, entre éstas, habrá algunas acciones deseables y otras no. Sin embargo, una intención, en tanto que se dirige al futuro, no puede distinguir una acción particular en un sentido inteligible. Pienso que la perplejidad surge porque hemos pasado por alto una distinción importante. Sería absurdo sostener que cualquier acción mía en el futuro inmediato que sea comer algo dulce sería deseable. Pero no hay nada absurdo en que yo juzgue que cualquier acción mía en el futuro inmediato que sea comer algo dulce sería deseable *dado el resto de lo que creo acerca del futuro inmediato*. No creo que comeré un dulce venenoso y, así, ésa no es una de las acciones de comer algo dulce que mi juicio global incluye. Sería erróneo intentar mejorar el enunciado de mi intención diciendo «tengo la intención de comer algo dulce siempre que no sea venenoso». Como vimos, éste es un error porque si éste es el camino que debo recorrer, nunca obtendré correctamente mis intenciones. Hay un *sinsín* de circunstancias conforme a las cuales yo no comería algo dulce y no puedo comenzar a preverlas todas ellas. La cuestión es que no creo que nada surgirá que haga indeseable o imposible mi comer. Esa creencia no es parte de aquello de lo que tengo la intención de hacer, sino que es un supuesto sin el cual yo no tendría la intención. La intención no es de forma condicional sino, más bien, la existencia de la intención está condicionada por mis creencias.

Tengo la intención de comer mañana un desayuno sustancioso. Tú y yo sabemos que mañana no comeré un desayuno sustancioso si no tengo hambre. Y no estoy seguro de que estaré hambriento, tan sólo pienso que lo estaré. En estas condiciones, no sólo no es más preciso decir «tengo la intención de comer un desayuno sustancioso si estoy hambriento»: es *menos* preciso. Yo tengo la segunda intención, tanto como la primera, pero la primera implica la segunda y no a la inversa y, así, la primera es una versión más completa de mis intenciones. Si tú tan sólo supieses que yo tenía la intención de comer un desayuno sustancioso si yo estuviese hambriento, no sabrías que creo que estaré hambriento, que es lo que sucede realmente. Pero tú podrías colegir esto si supieses que tengo la intención de comer un desayuno sustancioso mañana.

Creo que esta visión del asunto explica el problema que tuvimos acerca de la relación entre tener la intención de hacer algo y creer que uno lo hará —por qué, por una parte, es tan extraño decir «tengo la intención de hacerlo, pero quizás no lo haré» y, sin embargo, por otra parte, es tan imposible aumentar la precisión de los enunciados de intención condicionalizando el contenido de la intención con respecto a cómo resulten las cosas. La explicación es que la intención supone, pero no contiene, una referencia a cierta visión del futuro. Una intención actual con respecto al futuro es, en sí misma, como un informe provisional: dado lo que ahora sé y creo, he aquí mi juicio de qué tipo de acción es deseable. Puesto que la intención se basa en la mejor apreciación que uno tiene de la situación, tan sólo distorsiona las cosas decir que el agente tiene la intención de actuar de la manera como lo hace sólo si su apreciación resulta ser correcta. Una intención actual no necesita ser algo similar a una resolución o a un compromiso (aun cuando *decir* que uno tiene la intención de hacer algo puede, en ocasiones, tener este carácter). Mi intención se basa en mi visión actual de la situación; no hay ninguna razón, en general, por la que yo haya de actuar como ahora tengo la intención de hacerlo, si mi visión actual resulta ser equivocada.

Ahora podemos ver por qué añadir «si puedo» nunca hace más precisa la formulación de una intención, aun cuando puede servir para cancelar una sugerencia natural no deseada del acto de decir que uno tiene la intención de hacer algo. Tener la intención de realizar una acción es, conforme a mi versión, sostener que es deseable

realizar una acción de cierto tipo a la luz de lo que uno cree que pasa y de lo que pasará. Pero si uno cree que no es posible ninguna acción semejante, entonces no puede haber ningún juicio de que es deseable tal acción consistente con las propias creencias. No puede haber ninguna intención así.

Si una intención es tan sólo un juicio de que es deseable una acción de cierto tipo, ¿qué es lo que distingue una intención de un mero deseo? Podemos dejar de lado deseos de cosas que no son consistentes con lo que uno cree, pues éstas están excluidas por nuestra concepción de una intención. Y podemos dejar de lado deseos que no correspondan a juicios globales. («Deseo poder ir a Londres la semana próxima»: mi ir a Londres la semana próxima puede ser consistente con todo lo que creo, pero no con todo lo que quiero. Este deseo es ocioso porque se basa sólo en algunas de mis razones *prima facie*.) Pero una vez que hacemos de lado estos casos, no hay necesidad de distinguir las intenciones de los deseos. Pues un juicio de que algo que pienso que puedo hacer y que pienso que veo claro el camino para hacerlo, un juicio de que tal acción es deseable no sólo por una u otra razón, sino a la luz de todas mis razones; un juicio como éste no es un mero deseo. *Es* una intención. (Esto no es negar que hay casos límite.)

¿Qué tan bien hemos manejado el problema con el que comenzamos? El problema consistía, en efecto, en dar una versión de tener la intención (y de formar una intención) que se trenzara de manera satisfactoria con nuestra versión de actuar con una intención y que no sacrificara los méritos de esa versión. Con respecto a la primera propuesta, encontrar una versión de tener la intención que se trenzara con nuestra versión de acción intencional, ideamos una manera satisfactoria de relacionar los dos conceptos, pero sólo mediante la introducción de un nuevo elemento en el análisis de la acción intencional: un juicio global. Dado este tipo de juicio y la idea de tal juicio formulado a la luz de lo que se cree acerca del acontecer futuro, creo que fuimos capaces de llegar a una tesis plausible de tener la intención.

Aún queda la cuestión acerca de si el tipo de juicio al que me he referido, un juicio global, puede entenderse sin apelar a las nociones de intención o voluntad. Pregunté, al principio de esta última sección de mi escrito, si tener la intención de hacer algo es querer hacerlo;

si lo fuese, podríamos considerar que hemos alcanzado nuestra meta. Lo que tenemos la intención de hacer queremos hacerlo, en un sentido muy amplio de querer. Pero esto no significa que tener la intención sea una forma de querer, pues considérense las acciones que quiero realizar y que son consistentes con lo que creo. Entre ellas están las acciones que tengo la intención de realizar y muchas más. Quiero ir a Londres la próxima semana, pero no tengo la intención de hacerlo, no porque piense que no puedo, sino porque interferiría con otras cosas que quiero más. Esto sugiere fuertemente que es mejor considerar querer y desear como si correspondieran a, o como si constituyeran juicios *prima facie*.

Si esto es correcto, no podemos sostener que hemos presentado argumentos convincentes a favor de ver las intenciones como algo familiar, como una especie de querer, donde podemos distinguir el tipo sin tener que usar el concepto de intención o el de voluntad. Sin embargo, lo que podemos decir es que tener la intención y querer pertenecen al mismo género de actitudes favorables expresadas por juicios de valor. El querer, los deseos, principios, prejuicios, deberes arraigados y obligaciones, proporcionan razones para las acciones y las intenciones y se expresan mediante juicios *prima facie*; las intenciones y los juicios que acompañan a las acciones intencionales se distinguen por su forma global o incondicional. Los casos de meramente tener la intención constituyen una subclase de los juicios globales, los que se dirigen a las acciones futuras del agente y que se hacen a la luz de sus creencias.