

MICHAEL SMITH

EL PROBLEMA MORAL

Traducción de
Rodrigo Sánchez Brigido
y Augusto Embrioni

Estudio preliminar de
Rodrigo Sánchez Brigido

Marcial Pons
MADRID | BARCELONA | BUENOS AIRES | SÃO PAULO
2015

moralmente buena es el hecho de que está motivada a hacer lo que cree que es correcto, cuando esto es leído *de dicto* y no *de re*. El externalismo entonces adopta una forma de fetichismo moral. Y segundo, vimos que el externalismo es incapaz de explicar por qué esperamos que las personas racionales hagan lo que les es requerido hacer. Para explicar esto necesitamos no sólo una concepción de los hechos morales que adopta el requerimiento de carácter práctico, sino también una concepción más fuerte de los hechos morales como hechos sobre requerimientos categóricos de la racionalidad o la razón.

Esto nos lleva a la tercera solución estándar al problema moral. Si pensamos que debemos aceptar tanto la afirmación de que nuestros juicios morales expresan nuestras creencias sobre una cuestión de hecho objetiva [esto es, 1)], y el carácter práctico del juicio moral [esto es, 2)], y pensamos que lo que esto muestra —quizás por las razones que se acaban de dar— es que debe haber hechos objetivos sobre las razones para la acción que todos compartimos, entonces podría pensarse que no tenemos más alternativa que rechazar la idea de HUME de que la creencia y el deseo son existencias distintas [esto es, 3)]. Ésta es la respuesta al problema moral favorecida por los teóricos antihumeanos de la motivación. Ellos piensan que la caracterización de HUME del modo en que la creencia y el deseo se combinan para producir la acción es un dogma en psicología filosófica, un dogma que no tenemos ninguna razón para aceptar.

Hemos considerado la teoría de la motivación antihumeana en detalle en este capítulo. Hemos visto que los antihumeanos fracasan en apreciar la distinción entre dos tipos de razones para la acción que tenemos: las razones motivacionales y las razones normativas. Y una vez que esta distinción está firme se vuelve claro que la teoría misma de la motivación del antihumeano enfrenta objeciones formidables. Pues sólo una teoría humeana de la motivación puede explicar adecuadamente por qué una explicación de las acciones de un agente en términos de sus razones motivacionales es ella misma una especie de explicación teleológica. La teoría de la motivación humeana pisa entonces suelo firme.

Las tres soluciones estándar al problema moral son por lo tanto todas *malas* soluciones. Necesitamos un enfoque diferente. Como yo lo veo, los antihumeanos estaban en lo cierto al pensar que la solución al problema moral descansa en el rechazo de una concepción completamente humeana de la psicología humana. Pero erraban al sostener que lo que debemos rechazar es la teoría de HUME de los distintos roles jugados por la creencia y el deseo en la motivación humana. Más bien, lo que debemos rechazar es la teoría característica de HUME de la racionalidad del deseo: su teoría de las razones normativas. Por lo tanto me dedico ahora a la tarea de proveer un análisis antihumeano de las razones normativas (capítulo V). Es aquí donde encontraremos nuestra solución al problema moral (capítulo VI).

CAPÍTULO V

UNA TEORÍA ANTIHUMEANA DE LAS RAZONES NORMATIVAS

1. DE LAS RAZONES MOTIVACIONALES A LAS RAZONES NORMATIVAS

El Humeano afirma que las razones motivacionales están constituidas por sus deseos y creencias medio-fin. He argumentado que tenemos buenas razones para aceptar esta afirmación. Sin embargo, una teoría de las razones motivacionales no agota una teoría psicológica, pues, como hemos visto, además de las razones motivacionales también hay razones normativas. Una teoría psicológica necesita por lo tanto decirnos algo sobre nuestras razones normativas que expresan normas de racionalidad o de razón también. Sobre lo que es y no es racional que nosotros hagamos.

La visión de HUME, por supuesto, es que mientras nuestras creencias pueden evaluarse en términos de verdad o falsedad, y por lo tanto pueden ser consideradas racionales o irracionales, nuestros deseos, que no pueden evaluarse en esos términos en absoluto, están fuera de la crítica racional por completo —al menos siempre que no estén basados en creencias falsas (capítulo I; SMITH, 1988b)—. La teoría humeana característica de las razones normativas, es, por lo tanto, que lo racional para un agente es simplemente actuar de manera que satisfaga de modo máximo sus deseos, cualquiera que sea el contenido de tales deseos. Lo que es racional para un agente hacer es por lo tanto relativo a lo que más quiere hacer. Ésta es la llamada concepción «maximizadora» de la racionalidad (GAUTHIER, 1975). La cuestión que emerge, naturalmente, es si debemos aceptar alguna versión de la teoría humeana de las razones normativas también.

A pesar de nuestra defensa anterior de la teoría humeana de las razones motivacionales, mi propia visión es que no deberíamos aceptar la teoría Humeana de las razones normativas. Deberíamos aceptar más bien una teoría radicalmente antihumeana. Mi tarea en este capítulo es decir por qué, y elaborar la versión de la teoría antihumeana de las razones normativas que favorezco. Tales teorías de las razones normativas son controvertidas, sin embargo, y por buenas razones. Como yo lo veo, parte de esta controversia se centra en un problema que ha sido pasado por alto en general sobre la explicación de la acción. Empiezo por traer a la luz ese problema.

2. LO INTENCIONAL Y LO DELIBERATIVO

Adviértase que podemos, en general, explicar la acción intencional desde dos perspectivas bastante diferentes: la *intencional* y la *deliberativa* (PETTIT y SMITH, 1990; PETTIT, 1993: 20-24). Aunque a primera vista esto parece no problemático, cuando reflexionamos aparece un problema. Permítanme explicar las dos perspectivas.

Desde la perspectiva intencional, explicamos la acción intencional haciéndolo encajar en un patrón de explicación teleológica, y quizás causal: en otras palabras, explicamos citando el complejo de estados psicológicos que producen la acción. Consideremos el hecho de que yo teclee estas palabras. Explicamos esta acción desde la perspectiva intencional cuando citamos mi deseo de escribir un libro y mi creencia de que puedo hacerlo tecleando estas palabras, pues éste es el par deseo/creencia que teleológicamente, y quizás causalmente, explica mi teclear estas palabras. En términos de nuestra distinción entre dos tipos de razones, explicamos citando mis *razones motivacionales*.

Desde la perspectiva deliberativa, por contraste, explicamos una acción intencional en términos del patrón de deliberación racional que, o bien la produjo, o bien pudo haberla producido. Considérese de nuevo mi teclear estas palabras. Al decidir si teclear estas palabras o no reflexiono sobre ciertos hechos: que sería deseable escribir un libro y que puedo hacerlo tecleando estas palabras. Estas consideraciones están dentro de las consideraciones que de hecho tengo en cuenta al decidir qué hacer antes de hacerlo. Dan mis razones. Constituyen mi justificación racional. En términos de nuestra distinción entre dos tipos de razones, estas consideraciones constituyen mis *razones normativas* para hacer lo que hago, al menos tal como esas razones se me aparecen. Por supuesto, sería erróneo suponer que atravesamos conscientemente dicho proceso de razonamiento cada vez que actuamos, pero, aun cuando no lo hacemos, con frecuencia podemos reconstruir el patrón

de razonamiento que podría haber sido explícito en nuestro decidir hacer lo que hicimos. En este caso, la justificación *ex post facto* toma la forma de la construcción de un relato «como si», un relato que puede estar más o menos cerca de la verdad.

A primera vista, entonces, la distinción entre lo intencional y lo deliberativo parece involucrar una diferencia de *categorías*. Pues, mientras que desde la perspectiva intencional estamos interesados en qué estados psicológicos del agente explican sus acciones, desde la perspectiva deliberativa estamos interesados en qué proposiciones, desde el punto de vista del agente, justifican sus acciones. Como tal, podría pensarse que no hay ningún conflicto al nivel de la explicación. Sin embargo, examinado el asunto más de cerca, vemos que ambas perspectivas nos comprometen con afirmaciones sobre la explicación que están en un conflicto potencial entre sí.

Para ver que la perspectiva deliberativa es de hecho una perspectiva sobre la explicación, es suficiente advertir que afirmar lo contrario es equivalente a suponer que la conexión entre lo que decidimos hacer, sobre la base de la deliberación racional, y lo que hacemos, es completamente contingente y fortuito. Y esto es patentemente absurdo. Cuando deliberamos, y decidimos qué justificación racional tenemos para actuar, ese mismo hecho a veces hace una diferencia a lo que hacemos. Pero si la deliberación es, al menos algunas veces, práctica no sólo en su contenido sino también en su resultado, entonces debemos suponer que nuestras actitudes hacia las proposiciones que figuran en nuestras deliberaciones son capaces de figurar en una explicación de lo que hacemos. Y ahora el potencial para un conflicto emerge.

Considéremos de nuevo nuestro ejemplo. Si creo que es deseable escribir un libro —o, de modo equivalente, si pienso que tengo una razón normativa para hacerlo— y creo que puedo escribir un libro tecleando estas palabras, entonces podemos redesccribir mis actitudes en estos términos: yo *valoro* escribir un libro y *creo* que puedo escribir un libro tecleando estas palabras. Podríamos decir, entonces, que en nuestra explicación deliberativa las actitudes en cuestión son valorar y creer. Pero ahora debemos preguntarnos cómo esta explicación deliberativa en términos de *valorar* y creer se relaciona con la explicación intencional de la misma acción en términos de *desear* y creer. La actitud problemática es la valoración. ¿Qué es valorar? ¿Y cómo el valorar se relaciona con el desear?

3. ALGUNAS DIFERENCIAS ENTRE VALORAR Y DESEAR

Queremos responder a la pregunta «¿qué es valorar?». Al responder a esta cuestión debemos permanecer fieles a nuestros pensamientos ordinarios

sobre las relaciones entre nuestras valoraciones, por un lado, y nuestros deseos por el otro. En otras palabras, debemos permanecer fieles a las varias trivialidades sobre la conexión o falta de conexión que encontramos entre lo que consideramos racionalmente justificado y lo que estamos motivados a hacer. Ya he dicho por qué debemos pensar que hay alguna conexión. En esta sección explico por qué también debemos pensar que hay alguna falta de conexión.

Considérese el siguiente pasaje del libro de AYER *Freedom and Necessity*.

El cleptómano no atraviesa ningún proceso de decidir si robar o no robar. O más bien, si atraviesa dicho proceso, es irrelevante para su conducta. Lo que quiera que haya resuelto hacer, de todos modos robará. Y esto es lo que lo distingue de un ladrón ordinario (AYER, 1954: 20).

Ayer correctamente considera como dado que el cleptómano roba intencionalmente. Esto es, que lo que hace es explicable en términos de lo que quiere hacer. El problema que él señala, sin embargo, es que si bien la acción del cleptómano es por lo tanto explicable desde la perspectiva intencional, no necesita ser explicable desde la perspectiva deliberativa. Pues hay, tal como nota AYER, un potencial para una brecha entre lo que el cleptómano «resuelve» hacer como resultado de la deliberación y lo que quiere hacer: una brecha en lo que él considera que tiene una razón normativa para hacer y lo que tiene una razón motivacional para hacer. Esto es importante, pues sugiere una brecha entre valorar y desear. En particular, sugiere que un agente puede desear actuar de cierto modo sin valorar actuar de ese modo: esto es, sin aceptar que tiene una justificación racional para actuar de ese modo.

Harry FRANKFURT formula un argumento similar. Él nos pide que imaginemos a un adicto a la heroína quien

Odia su adicción y siempre lucha desesperadamente aunque sin éxito en contra de su empuje. Él intenta hacer todo lo que cree que le permitirá superar sus deseos por la droga. Pero estos deseos son demasiado poderosos para que él los resista, e invariablemente, lo conquistan. Es un adicto contra su voluntad, violado por sus propios deseos sin esperanza (FRANKFURT, 1971: 87).

El adicto a la heroína ciertamente quiere tomar la droga. Sin embargo, tal como FRANKFURT nota, podemos imaginarlo diciendo que «“realmente” no quiere tomar la droga». O incluso que «preferiría morir» en vez de tomarla (1971: 83). Aquí, como en otros lugares, hablar de lo que «realmente quiere» es un sustituto para hablar de lo que valoramos o lo que creemos racionalmente justificable. El punto de FRANKFURT, como el de AYER, es en-

tonces que podemos desear actuar de cierto modo sin valorar actuar de ese modo. Sin pensar que actuar así está racionalmente justificado.

Gary WATSON formula un punto similar.

Considere el caso de una mujer que tiene un súbito deseo de ahogar a su bebé en la bañera; o el caso de un jugador de *squash* que, mientras sufre una derrota ignominiosa, desea golpear a su oponente en la cara con la raqueta. Es simplemente falso que la mujer valora el ahogamiento de su hijo o que el jugador valora el daño y el sufrimiento de su oponente. Pero sin embargo ellos desean estas cosas. Ellos desean a pesar de sí mismos. No es que asigne a esas acciones un valor inicial que después es sopesado contra otras consideraciones. Estas actividades ni siquiera son representadas con una entrada positiva, ni siquiera pequeña, en la «matriz inicial de la deseabilidad» (WATSON, 1975: 101).

Así, si la mujer que WATSON describe de hecho ahoga su bebé entonces, aunque ella puede actuar intencionalmente, no actúa de un modo que ella valora. Actúa sin pensar que lo que hace está racionalmente justificado en absoluto. Lo mismo se aplica al jugador de *squash* derrotado y enojado también.

Lo que los casos de AYER-FRANKFURT-WATSON muestran, entonces, es que podemos desear algo sin valorarlo. Sin pensar que, si actuamos en función de nuestros deseos, actuamos de modo racionalmente justificado. Pero hay otros casos también, casos que sugieren que podemos pensar que estamos racionalmente justificados al actuar de cierto modo sin desear actuar de esa manera.

Como vimos antes, Michael STOCKER observa que varios factores emocionales y físicos — «cansancio espiritual o físico», «indolencia», «debilidad del cuerpo», «enfermedad», «apatía general o desesperanza», «inhabilidad para concentrarse», «un sentimiento de inutilidad o futilidad», etc.; «depresiones», como en general él las llama— pueden socavar nuestro deseo de hacer lo que valoramos hacer (capítulo IV; STOCKER, 1979: 744). Y lo que STOCKER dice aquí sobre la brecha entre nuestros juicios acerca de lo que es bueno y lo que deseamos se aplica igualmente a casos en los cuales lo que está en juego es qué justificación racional tenemos para actuar. En estos casos también varias «depresiones» pueden socavar nuestro deseo por completo. El depresivo puede así conocer perfectamente bien que lo racional para él, por ejemplo, es levantarse y seguir con su vida: ir al trabajo, visitar un amigo, leer un libro, cocinar una comida o lo que fuese. Pero el efecto de su depresión puede ser precisamente remover cualquier deseo que tenga de hacer cualquiera de estas cosas. Al no tener ningún deseo de levantarse y seguir con su vida, puede por lo tanto simplemente hacer nada —o, en cualquier caso, nada intencionalmente—.

La significación de estos casos debería ser clara. Nos recuerdan no sólo que nuestras razones motivacionales pueden separarse de las afirmaciones de razones normativas que aceptamos, sino también que este tipo de separación ocurre por cierto tipo característico de razones. Así los casos AYER-FRANKFURT-WATSON nos recuerdan que, como resultados de compulsiones psicológicas, adicciones físicas y disturbios emocionales, podemos tener una razón motivacional para actuar de cierto modo sin pensar que podemos proveer una justificación racional para actuar de ese modo. Y los casos de STOCKER nos recuerdan que, como resultado de cansancio espiritual, indolencia, enfermedad, y cosas por el estilo, podemos carecer de ninguna razón motivacional para hacer lo que pensamos que tenemos una justificación racional para hacer. STOCKER mismo resume el punto de este modo.

Por supuesto, citar el bien (en que alguien cree) puede siempre ser una razón en el sentido de ser una razón justificatoria. Pero eso sólo significa decir que lo que sirve como una razón justificatoria puede no servir para hacer un acto inteligible, y lo que puede servir para hacer un acto inteligible puede no ser una razón justificatoria, sino una razón «desjustificatoria». (STOCKER, 1979: 746).

Los casos en que ocurren separaciones de este tipo entre las afirmaciones de razón normativa que aceptamos y nuestras razones motivacionales son todos casos desviados, por cierto. Pero el pensamiento ordinario nos dice que tales casos existen y que por lo tanto no deben ignorarse cuando proponemos una explicación de la relación entre valorar o aceptar razones normativas, por un lado, y desear, por el otro.

4. EL PROBLEMA

He dicho que el hecho de que la acción intencional pueda explicarse tanto desde la perspectiva intencional como desde la deliberativa genera un problema. Ahora estamos en posición de explicar cuál es el problema exactamente.

En la medida en que la reflexión sobre nuestras razones normativas nos mueve a actuar —esto es, en la medida en que somos deliberadores eficaces — aceptar que tenemos ciertas razones normativas debe de algún modo estar vinculado con tener las razones motivacionales correspondientes. Pero los casos desviados recién descritos nos recuerdan que nuestros deseos pueden estar separados de nuestra aceptación de las afirmaciones de razones normativas correspondientes, y que por lo tanto podemos ser deliberadores ineficaces. El problema, entonces, es explicar cómo puede ser que aceptar

afirmaciones de razones normativas puede a la vez *estar vinculado con tener deseos* y sin embargo *estar desvinculado de tener deseos*. En otras palabras, el problema es explicar cómo la deliberación sobre la base de nuestros valores puede ser práctica en su resultado *justo en la medida en que lo es*.

El problema aquí es profundo, un problema cuyo origen puede ser rastreado en la psicología basada en el par creencia-deseo que hemos heredado de HUME (capítulo IV). Para suponer que la deliberación es práctica en su resultado debemos suponer que hacemos lo que hacemos porque valoramos lo que valoramos. Y para suponer que la deliberación fracasa en ser práctica en su resultado debemos suponer que el hecho de que valoramos lo que valoramos es irrelevante para hacer lo que hacemos. ¿Pero qué es valorar algo? Esto es, dicho de otro modo, ¿qué es aceptar que tenemos una razón normativa para hacer algo? ¿En términos de HUME, es acaso una cuestión de *creencias*? ¿O es una cuestión de *deseo*? Parece que enfrentamos un dilema.

Si valorar es una cuestión de creer entonces tenemos una explicación fácil de por qué nuestros valores —considerados ahora como creencias sobre nuestras razones normativas— pueden estar separados de nuestros deseos. Pues las creencias y los deseos son existencias distintas. Pero por esas mismas razones entonces es difícil ver cómo lo que quiera que hagamos intencionalmente podría ser hecho *porque* valoramos lo que valoramos. Si nuestras creencias no pueden producir acciones —si, más bien, son simplemente representaciones inertes de cómo son las cosas— ¿entonces cómo nuestras creencias sobre las razones normativas pueden alguna vez hacer una diferencia respecto de lo que hacemos? Y si valorar es una cuestión de desear entonces nuestros problemas son exactamente los opuestos. Pues aunque será entonces claro cómo podemos actuar debido a nuestros valores, dada la visión humeana de la motivación es difícil ver cómo puede haber la brecha requerida entre lo que valoramos y lo que estamos motivados a hacer. Si aceptar lo que tenemos razón normativa para hacer es simplemente desearlo, ¿entonces cómo podemos fracasar en estar motivados a hacer lo que aceptamos que tenemos una razón normativa para hacer?

Lo que está en juego es entonces la coherencia misma de la idea de que la deliberación en función de nuestros valores es práctica en su resultado justo en la medida en que lo es. Como es quizás evidente, éste es el problema moral de nuevo, redescrito de modo de hacer explícito que lo que está en juego es la conexión entre la razón y la motivación. Y aquí, en consecuencia, es donde encontramos el corazón del debate.

En el resto de este capítulo considero las dos únicas alternativas disponibles: que valorar es una cuestión de desear y que valorar es una cuestión de creer. Para anticiparlo, argumentaré que valorar es una cuestión de creer, no de desear. La solución al problema yace en encontrar el contenido correcto

para la creencia; una teoría sobre la proposición que así es creída. Equipados con un análisis plausible de nuestro concepto de una razón normativa, argumentaré que podemos de hecho ver por qué nuestras creencias sobre nuestras razones normativas están necesariamente conectadas con nuestros deseos justo en la medida en que lo están.

5. DAVIDSON SOBRE VALORAR COMO DESEAR

La sugerencia más directa es que valorar es desear: esto es, que los deseos son expresados apropiadamente como afirmaciones acerca de lo que tenemos una razón normativa para hacer. Tal reducción de valorar a desear es propuesta por Donald DAVIDSON:

La expresión natural del [...] deseo [...] [de un agente] [...] [digamos, el deseo de mejorar el gusto del guiso] [...] tiene [...] forma evaluativa; por ejemplo, «es deseable mejorar el gusto del guiso». Podemos suponer que diferentes pro-actitudes son expresadas con otras palabras evaluativas en vez de «deseable».

No hay ninguna prueba simple de que las oraciones evaluativas expresan deseos y otras proactitudes del modo en que la sentencia «la nieve es blanca» expresa la creencia de que la nieve es blanca. Pero las siguientes consideraciones quizás ayuden a mostrar lo que está involucrado. Si alguien que sabe castellano dice honestamente: «la nieve es blanca» entonces cree que la nieve es blanca. Si mi tesis es correcta, alguien que dice honestamente «es deseable que deje de fumar» tiene alguna pro-actitud hacia su dejar de fumar. Tiene alguna inclinación a hacerlo; de hecho lo hará si nada se le pone en el camino, si sabe cómo, y si no tiene valores o deseos en contrario. Dada esta suposición, es razonable generalizar: si los juicios de valor explícitos representan proactitudes, todas las proactitudes pueden ser expresadas por juicios de valor que son al menos implícitos (1978: 86).

Aunque la sugerencia de DAVIDSON es sencilla, está profundamente equivocada. Pues simplemente ignora la posibilidad de los casos desviados ya mencionados. Tampoco es difícil ver dónde ha cometido la equivocación DAVIDSON. En pocas palabras, asume incorrectamente que un rasgo de los evaluadores *racionales* es un rasgo de *todos* los evaluadores. Así, desde cualquier punto de vista, un evaluador racional que dice honestamente «es deseable que deje de fumar» tiene de hecho alguna proactitud hacia dejar de fumar. Pero no se sigue de ello que alguien que es capaz de decir, honestamente, «es deseable que deje de fumar» pueda no tener una inclinación a dejar de fumar. Y tampoco se sigue que alguien que es incapaz de decir, honestamente, «es deseable que yo fume» pueda tener alguna inclinación a fumar. Todo lo que se sigue, si alguna de estas alternativas es posible, es

que los agentes en cuestión no son evaluadores racionales. Son depresivos o compulsivos o lo que quiera que su depresión o compulsión haya provocado que deseen, irracionalmente, algo distinto de aquello que creen que tienen una razón normativa para hacer.

Si la reducción de DAVIDSON de valorar a desear ha de ser exitosa, entonces, necesita convencer de que estas patologías no tienen tales consecuencias. Esto es, necesita argumentar que la descripción previa del pensamiento ordinario sobre las diferencias entre valorar y desear fueron *malas* descripciones. Que las razones motivacionales deben —por alguna razón que aún debe ser explicada— traer consigo la aceptación de afirmaciones de razones normativas correspondientes aun cuando los agentes tienen compulsiones, o están deprimidos, o lo que fuere.

De hecho, DAVIDSON provee tal argumento en su teoría de la racionalización.

A la luz de una razón primaria [...] el agente es visto en su rol de Animal Racional. De modo correspondiente a la creencia y a la actitud de una razón primaria para la acción [...] [esto es su causa] [...], podemos siempre construir (con un poco de ingenio) las premisas de un silogismo del que se sigue que la acción tiene (como la llama Anscombe) una «característica de deseabilidad». Así hay un cierto sentido irreductible —aunque algo anémico— en que cualquier racionalización justifica: desde el punto de vista del agente había, cuando actuó, algo para ser dicho en favor de la acción (1963: 9).

Así, de acuerdo con DAVIDSON, cuando racionalizamos una acción damos la razón de por qué fue hecha y así vemos al agente en su rol de «Animal Racional». Y lo que él de hecho argumenta es que esto es posible cuando citamos el par deseo-creencia que causa la acción sólo porque el deseo que causa la acción está apropiadamente expresado en una afirmación correspondiente de razón normativa. Si no fuera así expresado —DAVIDSON parece estar diciendo— no seríamos capaces de ver al agente en su rol de Animal Racional.

Michael Woods tiene una idea similar. Recuérdese su sugerencia de que «el concepto de una razón para la acción está en el punto de intersección [...] entre la teoría de la explicación de las acciones y la teoría de su justificación». Si dejamos de lado la dimensión explicativa, de acuerdo con Woods, no tenemos una razón para la acción. Pues nuestra apreciación de nuestras razones para la acción hace una diferencia a lo que hacemos. Pero si dejamos de lado la dimensión justificatoria, tal como lo ve, no tenemos una razón para la acción tampoco. Porque lo que distingue una explicación de una acción en términos de las razones para la acción del agente de otros tipos de explicaciones es el hecho de que una explicación en términos de razones no sólo explica sino que también justifica.

La reducción de DAVIDSON de valorar a desear es por lo tanto un elemento crucial en una teoría de cómo las racionalizaciones pueden tener los rasgos que tanto él como WOODS creen que las racionalizaciones deben tener. Pues, si funciona, la reducción muestra cómo puede ser que los deseos expliquen y también justifiquen lo que hacemos. Los deseos, esos estados psicológicos que explican causalmente nuestras acciones, son apropiadamente expresados en evaluaciones, afirmaciones que nos permiten justificar nuestras acciones.

Sin embargo, a pesar del carácter ingenioso de esta sugerencia, me parece que es relativamente fácil ver adónde se equivoca. Hay una ambigüedad en la idea de las razones para la acción de un agente. Y, dada esta ambigüedad, hay por lo tanto dos modos en que podríamos ver a un agente en su rol de Animal Racional. Si la mujer descrita por GARY WATSON ahoga a su bebé en la bañera, por ejemplo, entonces, en un sentido, podemos decir que tiene tanto una razón para actuar como que actúa por esa razón. Pues tiene una razón motivacional y esa razón figura en una explicación teleológica de lo que hace. En la medida en que su razón motivacional figura en tal explicación vemos por lo tanto al agente en su rol de Animal Racional; pues lo que tenemos aquí no es una mera explicación causal de su conducta, sino más bien una explicación que nos permite verlo como persiguiendo un objetivo que tiene; una explicación que es suficiente para verlo como actuando intencionalmente (WILSON, 1985).

Pero en otro sentido, por supuesto, la mujer no tiene ninguna razón para actuar y lo que hace no es por lo tanto explicable, aun desde su punto de vista, como algo hecho por una razón. Pues ella reconoce que no hay ninguna justificación racional para lo que hace. Reconoce que el objetivo que persigue es él mismo injustificado e irrazonable. Lo que hace en modo alguno responde a sus pensamientos acerca de lo que es racional para ella hacer. No responde a sus pensamientos sobre sus razones normativas. En esta medida no podemos considerarla como actuando por una razón en absoluto. No podemos verla en su rol de Animal Racional.

La reducción de DAVIDSON de valorar a desear pasa por alto el hecho de que hay estos dos modos diferentes en que podemos ver a un agente en su rol como Animal Racional. Él asume, incorrectamente, que ver a un agente en este rol requiere que lo veamos como habiendo hecho algo que es racionalmente justificable, al menos desde su propio punto de vista, cuando en realidad requiere, como mucho, que lo veamos como persiguiendo un objetivo que tiene. Pero si esto es correcto entonces no tenemos ninguna razón para aceptar la reducción de DAVIDSON de valorar a desear en absoluto. Pues tal reducción simplemente no está requerida por una teoría de cómo la racionalización revela al agente su rol de Animal Racional¹.

¹ Michael STOCKER presenta el asunto de esta manera: «Se sostiene a menudo que la bondad de algo o la creencia en su bondad —su ser racional, dado los valores y creencias del sujeto—

6. GAUTHIER SOBRE VALORAR COMO UN MODO DE DESEAR

David GAUTHIER sugiere que valorar no es desear *simpliciter*, sino un modo de desear (1986: 22-23).

La racionalidad práctica en el sentido más general se identifica con la maximización [...] un objetor podría acordar en identificar la racionalidad práctica con la maximización, pero insistir en que una medida de preferencia individual no es la cantidad apropiada para ser maximizada. Es racional maximizar el valor. La teoría de la acción racional implícitamente identifica el valor con [...] [una medida precisa de preferencia: *i. e.* con] [...] la utilidad, pero el objetor desafía esta identificación [...] [Él podría] [...] acordar que el valor es una medida, pero insistir en que no se miden preferencias brutas, las cuales podrían estar desinformadas, no basadas en la experiencia, o no consideradas apropiadamente. Aceptaremos esta visión en la medida en que concierne el modo en que uno tiene preferencias.

Así, de acuerdo con GAUTHIER, un agente valora un cierto resultado sólo si desea ese resultado de *algún modo*: esto es, sólo si su deseo pasa ciertos *test* de reflexión y experiencia (1986: 29-32). Suponga que yo prefiero el vino blanco al tinto, pero sin haber nunca probado un sorbo de vino tinto. Esta preferencia puede que no pase el *test* de la experiencia, pues si probara el vino tinto podría advertir que prefiero el vino tinto al blanco. O suponemos que tengo que decidir si tomar vino tinto o blanco, y que elijo el blanco, revelando así mi preferencia, pero sin haber considerado el asunto en absoluto. Esta preferencia puede que no pase el *test* de la reflexión. Pues, reflexionando, podría haber encontrado que prefiero el vino tinto al blanco.

Aunque estas exigencias sobre el modo de desear apropiadas para el valorar no son, tal como han sido descritas, suficientes para descartar los casos desviados que hemos descrito, GAUTHIER podría evidentemente enriquecer su concepción del modo apropiado. Así podría agregar: «valorar es

hace inteligible (explica) por qué una persona lo busca o lo desea. Si lo que yo he dicho [...] es correcto, pues entonces eso es un error [...] Del mismo modo en que la persona bien puede no buscar o desear el bien (en el que cree), así, si tal persona fuera a hacer lo que produciría el bien (en el que cree), ese hecho podría no hacer inteligible por qué actuó de ese modo. Si se sabe que estoy hundido en una profunda desesperación o en algún otro tipo de depresión o que hace ya tiempo que he cesado de velar por el bienestar de alguien, entonces citar la bondad de mi acto (o la bondad en la que creo) no hará inteligible mi acto que beneficia a esa otra persona» (STOCKER, 1979: 746). Y nuevamente a continuación: «Así, la racionalidad en el sentido de maximizar el valor contra el trasfondo de las creencias de un agente no es la forma de toda acción, ni siquiera de toda acción inteligible. Ni es la persona racional correspondiente a la forma de todas las personas, ni siquiera de todas las personas inteligibles. Intentar comprender a la gente como si fuesen tales seres racionales implica psicologías morales inadecuadas» (STOCKER, 1979: nota 12).

un modo de desear donde el modo en cuestión descarta deseos que requieren apoyo de una compulsión psicológica, una adicción física, o un estado de perturbación emocional». De este modo él podría acordar con el pensamiento ordinario que podemos desear un resultado sin valorarlo.

Sin embargo, aunque estas modificaciones serían suficientes para hacer eso, fracasarían completamente en mostrar cómo podemos valorar un resultado sin desearlo. De hecho, si valorar es en algún sentido un *modo* de desear, como GAUTHIER sugiere, entonces valorar sin desear se convierte en una imposibilidad conceptual. Pero como hemos visto, el pensamiento ordinario nos dice que valorar sin desear no sólo es posible, sino que ocurre. Las «depressiones» que describe Michael STOCKER socavan nuestros deseos pero dejando intactas nuestras perspectivas evaluativas.

El problema con cualquier punto de vista de acuerdo con el cual valorar es sólo un *modo* de desear, entonces, es que sólo podrá dar cuenta del hecho de que podemos desear hacer cosas sin pensar que hacerlas está racionalmente justificado. Descartará la posibilidad misma de pensarnos racionalmente justificados al actuar de ciertos modos sin tener los deseos correspondientes. Pero dado que esos deseos son evidentemente posibles, se sigue que deberíamos rechazar la visión de que valorar es un modo de desear.

7. LEWIS SOBRE VALORAR COMO DESEAR DESEAR

Podría pensarse que una reducción de valorar a desear todavía es posible; que simplemente necesitamos reconfigurar el deseo en cuestión de un modo más radical. Por ejemplo, tanto Harry FRANKFURT (1971) como David LEWIS (1989) sugieren que podemos reducir valorar a *desear desear*. Aquí está LEWIS.

De modo que consideramos los deseos. Pero más vale que no digamos que valorar algo es exactamente lo mismo que desearlo. Eso podría ser el caso para algunos de nosotros: aquellos que logramos, por fuerza de voluntad o suerte, desear exactamente lo que deseamos desear. Pero no todos nosotros somos tan afortunados. El adicto pensante puede desear su aturdimiento eufórico, pero no valorarlo. Aun poniendo a un costado todos los costos y riesgos, puede odiarse a sí mismo por desear algo que no valora en absoluto. Es un deseo que quiere sacarse de encima. Él desea su euforia, pero no desea desearla. No desea un estado de conciencia mundano y no alterado, pero sí desea desearlo. Concluimos que no valora lo que desea, sino más bien que valora lo que desea desear (1989: 115).

La reducción de LEWIS de valorar a desear desear evita los problemas que enfrentaba DAVIDSON. Pues un agente ciertamente puede tener ciertos deseos

de primer orden distintos de sus deseos de segundo orden. Y esta reducción evita el problema que enfrentaba GAUTHIER también. Pues un agente puede tener un deseo de segundo orden distinto de su deseo de primer orden.

Podría pensarse entonces que la reducción de LEWIS nos da justo la distinción que queremos, la distinción entre la explicabilidad teleológica y la justificabilidad racional. La idea sería, ahora, que podemos decir que una acción es teleológicamente explicable si es explicable en términos de un deseo de primer orden, pero que para ser racionalmente justificable esa acción debe además estar de acuerdo con un deseo extra de segundo orden.

LEWIS se resiste a decir esto, sin embargo. Él niega que la acción *necesite* el respaldo de un deseo de segundo orden para ser racionalmente justificable (LEWIS, 1989: 116). Me parece que está en lo cierto al negar eso. Permítanme explicar brevemente por qué.

El punto que salió a la luz en nuestra discusión de las reducciones de DAVIDSON y GAUTHIER de valorar a desear fue que, aunque un agente racional desea —deseos de primer orden— de acuerdo a las afirmaciones de razones normativas que acepta, un agente irracional puede desear de manera diferente. Esto sugiere que las razones normativas están sujetas a la siguiente exigencia.

C1 Si un agente acepta que tiene una razón normativa para hacer φ entonces debería racionalmente desear φ .

C1 nos debería parecer familiar. Es similar al requerimiento de carácter práctico del juicio moral. Similar, esto es, a la afirmación de que si un agente juzga correcto hacer φ entonces, ausente alguna forma de irracionalidad práctica, está motivado a hacer φ (capítulo I). Si, como ya he argumentado, nuestro concepto de un requerimiento moral es él mismo simplemente el concepto de una razón normativa, entonces esta similitud no debería sorprendernos (capítulo III).

En una comprensión intuitiva, C1 le da a los agentes un consejo relativamente claro. Les dice que, permaneciendo todo lo demás igual, alguien que acepta que tiene una razón normativa para hacer φ y sin embargo, digamos, desea no hacer φ , debería sacarse de encima su deseo de no hacer φ y adquirir el deseo de hacer φ en cambio. El adicto a la heroína, por ejemplo, debería sacarse de encima su deseo por un aturdimiento eufórico, y adquirir un deseo por un estado de conciencia más mundano y no alterado.

Pero ahora consideremos C1 en el contexto de la reducción propuesta de valorar a desear desear. Si aceptar una afirmación de razón normativa es lo mismo que valorar, y valorar es simplemente desear desear, entonces se

sigue que un agente que desea desear hacer φ racionalmente debería desear hacer φ . ¿Pero es esto correcto? ¿Están nuestros deseos de primer y segundo orden relacionados normativamente de este modo? Considere de nuevo C1 en la comprensión intuitiva recién descrita. ¿Debería un agente que desea desear hacer φ , y que sin embargo desea no hacer φ , deshacerse de su deseo de no hacer φ y adquirir el deseo de hacer φ en cambio?

No necesariamente —o más bien, no necesariamente si asumimos, con el humeano, que lo que es racional hacer para alguien es una función del contenido y fuerza de sus varios deseos—. Esto es, no si asumimos, como en la concepción maximizadora de la racionalidad, que un agente debería hacer racionalmente lo que más quiere hacer. Pues todo depende de cuál de sus dos deseos es el más fuerte. Si su deseo de no hacer φ es más fuerte que su deseo de hacer φ , entonces lo racional para él es actuar en función de su deseo de no hacer φ y frustrar su deseo de hacer φ en cambio. Así, por ejemplo, si el deseo del heroinómano por su aturdimiento eufórico es más fuerte que su deseo de desear un estado de conciencia mundano y no alterado, entonces, en la concepción maximizadora humeana de la racionalidad, lo que él debería hacer es actuar de modo de satisfacer su deseo por el aturdimiento eufórico y frustrar su deseo de desear un estado de conciencia mundano. La razón está del lado del primero, no del segundo, violando C1.

Éstas son malas noticias para la reducción de valorar a desear desear. Y la mala noticia está, por supuesto, directamente relacionada al hecho de que la reducción es un análisis de valorar en términos de desear más que de creer. Pues aun si asumimos que la razón está del lado de una *armonía* entre nuestros deseos de primer y segundo orden (aunque vea SMITH, 1992), simplemente no hay una razón para asumir que la razón está del lado de lograr la armonía cambiando nuestros deseos de primer orden para que encajen con los deseos de segundo orden y no a la inversa. En la concepción humeana maximizadora de la racionalidad todo depende de qué deseo sea el más fuerte. Este tipo de problema no aparece para los análisis de valorar en términos de creer. Pues entonces, si hay algún cambio que deba hacerse, hay una razón de principio de por qué en los casos de desarmonía los deseos de un agente deben cambiar para acomodarse a sus valores: a saber, el hecho de que un deseo de primer orden en contrario del agente no será en modo alguno una señal del hecho de que su creencia evaluativa es falsa. Volveremos a este punto más adelante.

Podría pensarse que lo que esto muestra es simplemente que necesitamos modificar la concepción de la racionalidad maximizadora humeana. Después de todo, de acuerdo con esa concepción, los deseos tienen dos rasgos cruciales: contenido y fuerza. Quizás todo lo que necesitamos hacer es modificar la concepción maximizadora para dar cuenta de la significación

del contenido de un deseo de segundo orden. Quizás deberíamos agregar a la concepción maximizadora el siguiente principio:

- D Si un agente desea desear hacer φ entonces debería racionalmente desear hacer φ , y si desea desear hacer φ y desea no hacer φ entonces racionalmente debería deshacerse de su deseo de no hacer φ y adquirir en cambio el deseo de hacer φ .

En otras palabras, quizás deberíamos modificar la concepción maximizadora como para reconocer el hecho de que es importante si nuestros deseos son de primer o segundo orden. Desafortunadamente, sin embargo, esta respuesta yerra por completo.

Se supone que C1 actúa como una exigencia para una teoría adecuada de las razones normativas. Si podemos reducir la aceptación de afirmaciones de razones normativas a desear desear, entonces la esperanza es que esta reducción, en conjunción con otras suposiciones plausibles, efectivamente implicará este principio. Pero D es difícilmente una suposición adicional plausible. Es un principio teóricamente motivado que deberíamos aceptar sólo si se nos da un argumento adecuado. Pero el único argumento que se nos ha dado es que debe ser cierto si la reducción de valorar a desear desear es correcta. Y, en este contexto, ése no es un buen argumento.

El otro modo de expresar el mismo punto es el siguiente: la reducción de valorar a desear desear en conjunción con D captura de hecho el espíritu de C1. Pero la reducción misma no juega ningún rol significativo en ello. El espíritu de C1 es capturado por la conjunción de la reducción aun más *implausible* con un principio, como D, que estipula la conexión misma que queremos derivar. Las cosas serían diferentes, por supuesto, si tuviéramos una razón independiente para aceptar la reducción. Pero, en ausencia de esas razones independientes, simplemente no tenemos razón para adoptar la reducción aceptando D.

De hecho, sin embargo, ahora quiero argumentar que las cosas son todavía peores para la reducción de valor a desear de segundo orden. Pues no sólo no hay razones independientes para aceptar la reducción de valorar a desear de segundo orden, sino también que tales reducciones enfrentan una objeción decisiva (WATSON, 1975: 107-109). Por lo tanto tenemos todas las razones para rechazar directamente la reducción.

Aquellos que buscan reducir valorar a un deseo de segundo nivel de algún tipo deben ser claros e identificar valorar con algún deseo de orden más alto *en algún nivel u otro particular*. LEWIS es claro en ese sentido. Él identifica valorar con desear de segundo orden. ¿Pero por qué identificar valorar con desear de segundo orden? ¿Por qué no tercer orden o cuarto orden o ...?

La pregunta es una difícil para aquellos que proponen la reducción. Pues cada una de esas identificaciones parece tan plausible como cualquier otra. Y si cada una es tan plausible como la otra entonces *todas* esas identificaciones son igualmente implausibles. Pues cualquier identificación requeriría una elección arbitraria entre niveles. Por lo tanto no se ha propuesto ninguna reducción plausible en absoluto.

Lewis enfrenta esta objeción honesta y directamente, pero su respuesta es menos que convincente. Nos dice que su razón para favorecer el segundo nivel sobre el primero es que «un adicto reflexivo puede desear su aturdimiento eufórico, pero no valorarlo» (1989: 115). Y nos dice que su razón para favorecer cierto nivel más que otro de otro orden en el que el agente desea es que «si vamos al nivel más alto, automáticamente descartamos el caso de alguien que desea valorar de modo diferente al que lo hace, y sin embargo este caso no es obviamente imposible» (1989: 116). Hasta aquí esta línea de razonamiento parece perfectamente sólida. Los agentes pueden efectivamente desear actuar de modos que consideran que no pueden justificar racionalmente, y pueden de hecho desear pensar que actos que ellos no consideran racionalmente justificados lo están. Desafortunadamente, sin embargo, no hay más premisas en el argumento de LEWIS. A partir de esas premisas él concluye que valorar es desear de *segundo* orden.

El problema no es sólo que hay una brecha entre las premisas y la conclusión de LEWIS —decir que un deseo no es de primer orden y no es de un orden más alto no implica que es de segundo orden, algo que vemos inmediatamente si imaginamos a alguien con cuatro niveles de deseo—; el problema es que su argumento demuestra cuán formidable es la objeción original a la que supuestamente debía responder. Pues su conclusión es simplemente arbitraria, dadas las premisas. Él podría haber elegido del mismo modo cualquier otro nivel distinto del primero o del más alto. Y si esto es cierto entonces se sigue que no podemos identificar valorar con desear de desear en ningún nivel.

8. VALORAR COMO CREER

Si no podemos reducir valorar a desear entonces no tenemos más alternativa que considerar reducir valorar a creer. LEWIS, como DAVIDSON, rechaza la idea de que valorar es creer. ¿Pero por qué rechaza esta idea? Él razona del siguiente modo.

¿Qué es valorar? Es algún tipo de estado mental, dirigido hacia aquello que es valorado. Puede ser un sentimiento, o una creencia, o un deseo [...] *¿Un sentimiento?* Evidentemente no, porque los sentimientos que tenemos

cuando valoramos las cosas son demasiado diversos. ¿Una creencia? [...] Si valorar simplemente significa tener una cierta creencia sobre ello, entonces parece que no hay ninguna razón conceptual por la cual valorar es una actitud favorable. Podríamos no haber favorecido las cosas que valoramos. Nos podríamos haber opuesto, o haber sido completamente indiferentes. De modo que nos quedan los deseos (1989: 116).

El argumento de LEWIS en contra de identificar valorar con creer depende entonces crucialmente de la idea de que hay alguna conexión conceptual entre valorar y desear. ¿Pero conceder esta conexión conceptual realmente precluye identificar valorar con creer valorable?

LEWIS parece pensarlo así, pero no está del todo claro por qué. Después de todo, tal como él señala, el adicto puede desear su aturdimiento eufórico, pero no valorarlo. Y puede valorar un estado de conciencia mundano e inalterado pero no desearlo. Y a estos ejemplos podríamos agregar el de la mujer de WATSON con el bebé que llora, el ejemplo del jugador de squash derrotado y rabioso, y el de los depresivos de STOCKER. En otras palabras, no es solamente una posibilidad conceptual, sino que *realmente ocurre*, que somos indiferentes, o nos oponemos, a lo que valoramos. Cualquiera que sea la naturaleza precisa de la conexión conceptual entre valorar y desear, entonces, no precluye de modo obvio la suerte de indiferencia u oposición a lo que valoramos que la identificación de valorar con creer valorable hace posible.

Todo depende entonces de cuál es exactamente la naturaleza de la conexión conceptual entre creer valorable y desear. La respuesta apoyada por la discusión hasta aquí es que la conexión conceptual es simplemente la conexión *derrotable* descrita en C1. Todo lo demás igual, racionalmente deberíamos desear lo que valoramos. Esto es, reformulando C1 ahora en términos de «creencia» en vez de la palabra más neutral «aceptación», la conexión que buscamos es la siguiente:

- C2 Si un agente cree que tiene una razón normativa para hacer ϕ , entonces debería racionalmente desear hacer ϕ .

Y ahora enfrentamos el verdadero problema. ¿Cómo hemos de demostrar la posibilidad de este tipo de conexión conceptual entre nuestras creencias y deseos? ¿No estaría en conflicto cualquier demostración de ese tipo con la visión humeana, en favor de la cual he argumentado en el capítulo anterior, según la cual el deseo y la creencia son existencias distintas? Desde mi punto de vista, para responder estas preguntas primero debemos encontrar un análisis plausible de nuestro concepto de razón normativa. ¿Pero podemos proveer tal análisis?

Ciertamente ha habido intentos. Consideremos, por ejemplo, la siguiente sugerencia de Mark JOHNSTON.

En lo que respecta a asegurar una conexión interna o conceptual entre el valor y la voluntad, al menos *esto* es verdadero: en la medida en que uno no tiene debilidad de la voluntad uno deseará [...] lo que uno juzga valioso. Al menos eso es parte de la definición de la debilidad de la voluntad. En lo que respecta a hacer la conexión entre juzgar valorable y desear [...] particularmente inteligible, esto me parece que se logra por la observación de que «valioso» y «merecedor de ser deseado» son casi sinónimos. Si juzgar que algo es valioso es casi lo mismo que juzgar que merece la pena desearlo, entonces es rápidamente inteligible por qué juzgar algo valioso debería llevar a desear (1989: 161).

El argumento de JOHNSTON podría ser convincente si fuéramos a aceptar la afirmación de la casi sinonimia. ¿Pero deberíamos? ¿Son «valioso» y «merecedor de ser deseado» casi sinónimos?

JOHNSTON está en lo cierto al no decir que son sinónimos *reales*. Pues mientras que si hacer ϕ es valioso entonces hacer ϕ merece la pena ser *hecho*, si hacer ϕ es algo que merece la pena ser deseado entonces hacer ϕ es algo que merece la pena *desearse*, pero «que merece la pena hacer» significa algo diferente a que «merece la pena ser deseado». Sin embargo, aunque no son sinónimos reales, él podría estar en lo cierto en el sentido de que son «casi sinónimos». Eso sería así si, por ejemplo, se siguiera del hecho de que merece la pena hacer algo que es que merece la pena desearlo, y si se siguiera del hecho de que algo merece ser deseado que merece la pena hacerlo. Pero aun esta afirmación más débil me parece bastante equivocada. Pues, como Derek PARFIT ha señalado, puede que no sea deseable que deseemos hacer lo que es deseable que hagamos, y puede no ser deseable que hagamos lo que es deseable que deseemos hacer (1984: parte 1). Consideremos el ejemplo de PARFIT.

La teoría del autointerés me dice que la deseabilidad de una acción o un deseo es una función de la contribución que la acción o el deseo hace a mi auto-interés de largo término. Así es deseable que yo *haga* sólo una cosa: promover mi autointerés de largo término. Sin embargo, como señala PARFIT, no se sigue que sea deseable que yo desee promover mi autointerés de largo término. De hecho, bien puede ser indeseable que yo desee promover mi auto-interés de largo término. Todo depende de si tener ese deseo es necesario para que yo tenga el conjunto de deseos cuya posesión contribuirá más a mi autointerés de largo término. Y ese deseo puede que no sea necesario. Pues mi autointerés de largo término puede ser mejor servido por mi desear actuar en favor de mis familiares y amigos, escribir libros, promover

la humanidad, etc., sin tener ninguna preocupación directa en absoluto por mi autointerés de largo término.

Si esto es correcto entonces se sigue que la teoría del autointerés me dice que es deseable que yo desee hacer lo que no es deseable que haga. Y también me dice que es deseable hacer lo que no es deseable que yo desee hacer. Independientemente de cuán improbable esto pueda parecer, el punto que quiero enfatizar es que la cuestión es una cuestión *sustantiva*, y por lo tanto requiere una respuesta *sustantiva*. La respuesta no debe ser determinada por un «casi» *fiat* conceptual.

La teoría del autointerés es, por supuesto, sólo un ejemplo. Pero lo que es verdadero de la teoría del autointerés puede ser verdadero de otras teorías sustantivas de la racionalidad práctica, aun de la teoría correcta. El argumento de JOHNSTON pasa por alto la posibilidad de este tipo de brecha entre lo que es deseable que hagamos y lo que es deseable que deseemos hacer. Como tal, falla en tanto intento de «hacer rápidamente inteligible por qué juzgar valorar debería llevar a desear».

Sin embargo, aunque el argumento de JOHNSTON descansa en una afirmación falsa de la «casi sinonimia» está, creo, en el camino correcto. Si algo puede hacer inteligible la conexión entre juzgar deseable y desear es un análisis de nuestro concepto de deseabilidad en términos de desear. Veamos entonces si podemos formular un argumento en esa línea de un modo que evite la objeción de PARFIT.

Hemos visto que decir que una acción es deseable es decir que tenemos una razón para hacerla, donde las normas relevantes de evaluación son las normas de la racionalidad. Pero ahora adviértase que podemos explicar ese concepto, el concepto de lo que tenemos una razón normativa para hacer, aunque de un modo algo diferente del sugerido por JOHNSTON. Pues es una trivialidad decir que lo que es deseable que hagamos es lo que desearíamos hacer si fuéramos completamente racionales. Que lo que tenemos una razón normativa para hacer es lo que desearíamos hacer si fuéramos completamente racionales (KORSGAARD, 1986). De manera que lo que necesitamos hacer es transformar esta trivialidad en un análisis.

En la sección siguiente nuestro cómo esto puede ser hecho. Si estoy en lo cierto esto será suficiente para mostrar que valorar es un asunto de creer. Y en la sección siguiente nuestro cómo ese análisis, en conjunción con otras suposiciones plausibles, es suficiente para hacer rápidamente inteligible por qué creer deseable debería llevar a desear: esto es, C2.

9. EL ANÁLISIS DE LAS RAZONES NORMATIVAS

De hecho la trivialidad está cerca de constituir ya un análisis. Simplemente tenemos que recordar por qué la trivialidad es una trivialidad, y después refinar nuestra comprensión de ella.

Como yo lo veo, la trivialidad está relacionada con todo un grupo de trivialidades sobre el *consejo*. Si usted no está seguro acerca de qué hacer en cierta situación, ¿cómo debería decidir qué hacer? La respuesta es que usted debería considerar la sabiduría de la gente; debería buscar consejo. Pero usted no debería preguntar a cualquier persona mayor de edad en busca de consejo. Debería preguntarle a alguien mejor situado que usted para saber qué hacer, alguien que lo conozca a usted bien. Con esta idea en el trasfondo, una interpretación natural de la trivialidad aparece. Pues, convenientemente idealizada, somos de hecho las mejores personas para darnos consejos a nosotros mismos. Permítaseme explicarlo.

La trivialidad nos dice que lo que es deseable para nosotros hacer es lo que desearíamos hacer si fuésemos completamente racionales. En otras palabras —y ahora realmente estamos transformando la trivialidad en un análisis, porque estamos haciendo explícitas distinciones que en el mejor de los casos sólo están implícitas en la trivialidad— nos dice que lo que es deseable para nosotros hacer en ciertas circunstancias —llamemos a estas circunstancias el «mundo posible evaluado»— es lo que, no tal como somos realmente, sino como seríamos en un mundo posible en que fuésemos completamente racionales —llamemos a esto el «mundo posible evaluador»— querríamos que nosotros hiciéramos en tales circunstancias. Esto es, nos dice que hechos sobre la deseabilidad de actuar de cierto modo en el mundo evaluado están constituidos por hechos sobre los deseos que tenemos *sobre* el mundo evaluado *en* el mundo evaluador.

Típicamente, por supuesto, el mundo evaluado será el mundo real. Así, lo que es deseable que nosotros hagamos en nuestras circunstancias reales es lo que nuestras contrapartes racionales, mirándonos tal como somos desde su perspectiva más privilegiada, querrían que nosotros hagamos en las circunstancias reales. En términos de la idea de trasfondo, los hechos sobre lo que es deseable que nosotros hagamos están constituidos por hechos acerca de lo que nos aconsejaríamos a nosotros mismos hacer si estuviéramos perfectamente posicionados para darnos un consejo a nosotros mismos.

Con esta idea acerca de cómo podemos transformar la trivialidad en un análisis como telón de fondo, quiero ahora refinar nuestra comprensión del análisis. Lo haré considerando un conjunto de preguntas acerca del análisis. Como yo lo veo, estas preguntas se centran en problemas clave que surgen

para cualquier análisis de las razones normativas, y nuestras respuestas a estas preguntas muestran por lo tanto cómo el análisis ofrecido da cuenta de esos problemas clave.

Primera pregunta. «El análisis nos dice que lo que es deseable que nosotros hagamos en ciertas circunstancias C es lo que desearíamos hacer en C si fuésemos completamente racionales. ¿Pero acaso el análisis, como el de Johnston, no confunde el hecho de que merece la pena hacer φ con el hecho de que merece la pena desear φ ?».

No, no lo hace. Supongamos, por mor del argumento, que lo que tenemos razón para hacer, en nuestras circunstancias actuales, es promover nuestro auto-interés de largo término. De acuerdo con el análisis esto significa que, si yo fuera completamente racional, desearía que, en mis circunstancias reales, yo promueva mi autointerés de largo término. Pero adviértase que esto es un deseo sobre lo que he de *hacer* en mis circunstancias reales. No es un deseo sobre lo que he de desear en mis circunstancias reales. Lo que vale la pena hacer es promover mi auto-interés de largo término.

Sin embargo, podríamos suponer, es consistente con el hecho de que yo tendría este deseo si fuese completamente racional que también, si fuese completamente racional, deseara que cuando me encuentre en mis circunstancias reales, yo desee actuar por el bien de mi familia y amigos, escribir libros, promover la humanidad, etc., sin tener ninguna preocupación directa por mi autointerés de largo término. Esto es lo que desearía *desear* en mis circunstancias reales. Y, de modo acorde, lo que vale la pena desear es actuar por el bien de mi familia y amigos, escribir libros, promover la humanidad, etcétera.

En resumen, entonces, el análisis nos dice que si hacer φ es algo que vale la pena hacer y algo que vale la pena desear, o es sólo una de las dos cosas, ello depende de si, si fuese completamente racional, lo que desearía *hacer* en ciertas circunstancias y lo que desearía *desear* en esas mismas circunstancias es lo mismo o algo diferente. Esta es una cuestión sustantiva, una cuestión que debe ser determinada por los deseos de la persona completamente racional, no una cuestión a ser decidida por un fiat conceptual.

Segunda pregunta. «De acuerdo con el análisis “hacer φ en las circunstancias C es lo que desearíamos hacer en C si fuésemos completamente racionales” da el contenido de la idea “hacer φ es deseable”. Pero si esto fuera cierto ¿no se seguiría que nuestras creencias sobre los valores son creencias sobre nuestros propios deseos? ¿Y no estaría ello bastante mal? Los juicios de valor no son afirmaciones introspectivas sobre nuestros deseos, son afirmaciones sobre un estándar contra el cual medir nuestros deseos inspeccionables».

La afirmación de que el análisis transforma los juicios de valor en juicios introspectivos sobre nuestros propios deseos está simplemente equivocada. Más bien, nos ofrece un contraste vívido entre los juicios introspectivos sobre nuestros propios deseos y nuestros juicios de valor, un contraste que da un estándar natural contra el cual nuestros deseos (que pueden ser inspeccionados), pueden ser medidos. Para ver por qué consideremos un ejemplo.

Supongamos que Ana tiene muchas ganas de bailar una jiga. Dándose cuenta de su deseo de bailar una jiga, se pregunta a sí misma si vale la pena tener ese deseo y actuar en función del mismo. En otras palabras, ella delibera. Cuando ella delibera, ¿cómo figura su deseo en la toma de su decisión, dado el análisis? El deseo figura en su toma de decisión en al menos dos modos. Primero, una vez inspeccionado, su deseo de bailar una jiga figura como un objeto de evaluación. Se pregunta a sí misma si tener ese deseo es deseable. En este caso, su deseo de bailar una danza figura como un objeto de evaluación. Se pregunta a sí misma si tener este deseo es deseable. En este caso, Ana de hecho se concentra en sí misma y en sus propios deseos. Sin embargo esto es poco sorprendente dado que, *inter alia*, es su deseo actual de bailar una jiga lo que está siendo evaluado. Segundo, no obstante, una representación de los deseos que Ana tendría si fuese completamente racional también figura en su toma de decisión. Pues Ana se pregunta a sí misma, en efecto, si ella desearía desear, si fuese completamente racional, bailar una jiga en sus circunstancias actuales, ¿pero es esto un juicio introspectivo? Ciertamente no. Pues lo que hace al juicio verdadero, si es verdadero, no es un hecho sobre Ana a partir de la inspección, es más bien un hecho hipotético sobre Ana: esto es, el hecho acerca de qué querría ella si fuese completamente racional.

Más aún, mientras que la afirmación que Ana puede hacer acerca de sí misma mediante la introspección —la afirmación de que tiene cierto deseo, digamos el deseo de bailar una jiga— no es una justificación en absoluto para suponer que éste es un deseo que vale la pena tener o en función del cual vale la pena actuar, la afirmación de que, si fuese racional, ella querría, en las circunstancias en que actualmente se encuentra tener el deseo de bailar una jiga y actuar según ese deseo, sí provee una justificación para suponer que su deseo es algo que vale la pena tener y a partir de lo cual vale la pena actuar. Pues constituye el fundamento del consejo que su contraparte más perfecta le daría a su contraparte menos perfecta sobre qué desear y hacer en sus circunstancias actuales. Este hecho hipotético sobre los deseos de Ana, entonces, ofrece un estándar bastante plausible en contra del cual medir sus deseos, los que pueden ser inspeccionados, tal como la objeción insiste.

Tercera pregunta. «¿Qué nos dice el análisis sobre los casos desviados mencionados antes: el cleptómano de AYER, el heroinómano de FRANKFURT,

la madre descrita por WATSON que desea ahogar su bebé en la bañera, y los depresivos de STOCKER?».

Lo que estos casos desviados muestran, desde cualquier punto de vista, es que las compulsiones psicológicas, las adicciones físicas, los disturbios emocionales, la depresión, el cansancio espiritual, la indolencia, la enfermedad, y cosas por el estilo, tienen el potencial de causar que deseemos hacer lo que creemos que no tenemos ninguna razón para hacer por un lado, y no desear lo que creemos que tenemos una razón para hacer por el otro. El análisis sugiere una explicación directa de estas brechas entre nuestras creencias sobre nuestras razones normativas y nuestros deseos. Pues, crucialmente, una cosa es qué crea un agente que desearía hacer en ciertas circunstancias si fuera completamente racional, y otra cosa completamente distinta es que realmente desee actuar de ese modo en esas circunstancias. Los varios factores mencionados —compulsiones psicológicas, adicciones físicas, disturbios emocionales y cosas por el estilo— pueden ser causalmente responsables por los deseos que un agente realmente tiene. Pero los efectos de estas afecciones sobre los deseos de un agente pueden bien ser desplazados cuando forma sus creencias acerca de qué querría si fuese completamente racional. Pues tales afecciones no tienen ningún rol constitutivo en la psicología de un agente completamente racional.

El punto no es que los agentes que sufren de tales afecciones son necesariamente irracionales: pueden serlo o no (STOCKER, 1979). El punto es más bien que, si los deseos que tales agentes tienen acerca de lo que han de hacer en distintas circunstancias —incluyendo, por supuesto, las circunstancias en que sufren de una u otra de esas afecciones (PETTIT y SMITH, 1993b)— no serían tenidos en la *ausencia* de tales afecciones, entonces esos deseos son irracionales. Los deseos son irracionales en la medida en que son *completa y únicamente* el producto de compulsiones psicológicas, adicciones físicas, perturbaciones emocionales y cosas por el estilo. En la medida en que no serían tenidos por alguien en un estado no depresivo, no adictivo o no perturbado emocionalmente.

Así, por ejemplo, la mujer que de hecho quiere ahogar a su bebé en la bañera puede desplazar imaginativamente los efectos de su furia y frustración sobre sus deseos al formar sus creencias sobre qué, en condiciones de racionalidad plena, desearía hacerle a su bebé cuando sufre de una perturbación emocional. Ésta es la razón por la que, aun cuando desea ahogar a su bebé, puede no creer que tiene ninguna razón normativa para hacerlo. Y el depresivo puede desplazar imaginativamente los efectos de su depresión sobre sus deseos al formar sus creencias acerca de qué, en condiciones de racionalidad plena, desearía hacer cuando sufra de una depresión paralizante. Ésta es la razón por la que, aun cuando no tiene un deseo de levantarse

y seguir con su vida, bien puede creer que tiene una razón normativa para hacerlo. El análisis nos da entonces una comprensión directa e intuitiva de los casos desviados.

Cuarta cuestión. «La idea de alguien plenamente racional juega un rol crucial en el análisis. ¿Pero en qué consiste exactamente esa idea? ¿Puede hacerse más explícita la idea?»

Como yo lo veo, la idea de alguien plenamente racional es ella misma una noción *sumaria*. El rol de esta idea en el análisis es entonces capturar, a modo de resumen, toda una serie de trivialidades específicas sobre la racionalidad práctica. En efecto, ya he mencionado algunas de ellas. Pues decir que la depresión, perturbaciones y cosas por el estilo no tienen un rol constitutivo en la psicología de un agente plenamente racional es, en efecto, decir que no hay trivialidades que conecten esas afecciones con nuestra idea de un agente racional.

Notése, sin embargo, que Bernard WILLIAMS hace la idea de un agente plenamente racional más explícita en su propio análisis, similar, de las razones (1980). Pues, de acuerdo con WILLIAMS, alguien tiene una razón para hacer ϕ en las circunstancias C si y sólo si desearía hacer ϕ en las circunstancias C si fuese completamente racional, donde para ser racional el agente debe satisfacer las siguientes tres condiciones:

- i) El agente no debe tener creencias falsas.
- ii) El agente tiene que tener todas las creencias relevantes verdaderas.
- iii) El agente debe deliberar correctamente.

WILLIAMS argumenta que las condiciones i) a iii) deben satisfacerse concentrándose en varios ejemplos, ejemplos que serán lo suficientemente familiares a partir de nuestra discusión anterior sobre las diferencias entre las razones normativas y motivacionales (capítulo IV). En lo que sigue explicaré brevemente la idea detrás de cada una de estas tres condiciones, y después abordaré algunas de las cosas que Williams dice en conexión con la condición iii). Mi desacuerdo con WILLIAMS sobre la condición iii) será importante luego.

WILLIAMS argumenta en favor de la condición i) del siguiente modo. Suponga que un agente quiere mezclar un líquido de cierta botella con agua tónica y tomarlo. Sin embargo, él tiene ese deseo sólo porque desea tomar un *gin-tonic* y cree que la botella contiene ginebra, cuando en realidad contiene gasolina. Como WILLIAMS señala, «es simplemente muy extraño decir que tiene una razón para tomar ese líquido, y es natural decir que no tiene ninguna razón para tomarlo, aunque piense que la tiene» (1980: 102). ¿Por qué? Porque no tendría ese deseo si fuese completamente racional: esto es, *si no tuviera creencias falsas*.

WILLIAMS argumenta después en favor de la condición ii) haciendo notar que un agente «puede ser ignorante de algún hecho tal que si lo conociera, en virtud de algún elemento en [...] [su conjunto de deseos] [...], estaría dispuesto a hacer ϕ : podemos decir que tiene una razón para hacer ϕ , aunque no lo sabe» (1980: 103). Así, por ejemplo, si un agente desea comprar un Picasso y, aunque no lo sabe, hay de hecho un Picasso a la venta muy barato en la tienda de artículos usados, entonces ordinariamente diríamos que tiene una razón para comprar algo de esa tienda. ¿Por qué? Porque desearía hacerlo si fuese completamente racional: esto es, *si tuviera todas las creencias verdaderas relevantes*.

Y, finalmente, WILLIAMS argumenta en favor de la condición iii) en los siguientes términos. Hasta ahora hemos concedido que los deseos y las creencias interactúan de modos que generan nuevos deseos. Pero por supuesto esto es una afirmación sustantiva acerca de qué es ser racional. Nuestros deseos y creencias sólo generan nuevos deseos si deliberamos y lo hacemos correctamente. Así, por ejemplo, generan nuevos deseos sólo si razonamos de acuerdo con el principio medio-fin, pues sólo así un deseo por un fin se convierte en un deseo por un medio.

Más aún, tal como señala WILLIAMS, el razonamiento medio-fin es sólo un modo entre otros de la deliberación racional. Otro ejemplo es

[...] el razonamiento práctico [...] que lleva a la conclusión de que uno tiene una razón para hacer ϕ porque hacer ϕ sería lo más conveniente y placentero desde el punto de vista económico para satisfacer algún elemento en [...] [el conjunto de nuestros deseos] [...] y esto por supuesto es controlado por otros elementos en [el conjunto de nuestros deseos], aunque no necesariamente de un modo claro y determinado [...] [y] [...] hay posibilidades mucho más amplias para la deliberación, tales como: pensar cómo la satisfacción de los elementos en [nuestro conjunto de deseos] [...] puede ser combinada: *e. g.* en función del tiempo; cuando hay algún conflicto irresoluble entre los elementos del [...] [conjunto de nuestros deseos] [...] considerando cuál tiene más peso [...] o, de nuevo, encontrando soluciones constitutivas, tales como decidir qué sería una tarde entretenida, concedido que uno quiera entretenerse (1980: 104).

Y él piensa que hay otras posibilidades, más radicales, para la deliberación también.

Más sutilmente, [...] [un agente] puede pensar que tiene una razón para promover algún desarrollo porque no ha ejercitado su imaginación lo suficiente acerca de qué pasaría si lo hiciera. En su razonamiento deliberativo sin ayuda, o entusiasmado por las persuasiones de otros, puede llegar a tener algún sentido más concreto sobre qué estaría involucrado, y perder su deseo

por ello, así como de modo positivo, la imaginación puede crear nuevas posibilidades o nuevos deseos (1980: 104-105).

Así, de acuerdo con WILLIAMS, la operación de la imaginación también debe contar como un proceso deliberativo racional. Dada la amplia variedad de principios que gobiernan entonces la deliberación racional, él concluye que un agente tiene una razón para hacer ϕ sólo si desearía hacer ϕ si fuese completamente racional: esto es, *si delibera correctamente*.

En términos generales, las condiciones i) a iii) me parece que constituyen una elaboración bastante precisa de nuestra idea de la racionalidad práctica. Creo que requieren suplementarse, sin embargo. Por lo pronto, no veo ningún modo en que los efectos de las compulsiones, las adicciones, los disturbios emocionales, y cosas por el estilo podrían ser precluidas por las condiciones i) a iii) —a menos que alguna exigencia sea presupuesta en iii)—: la condición de la deliberación correcta. Además —y ahora más seriamente— me parece que WILLIAMS omite de su discusión de la condición iii) una teoría de lo que quizás sea la forma más importante de la deliberación. Dadas las conclusiones que quiere obtener de su análisis de las razones, conclusiones que discutiré ahora, ésta es una omisión seria. Permítaseme explicarlo.

WILLIAMS admite que la deliberación puede producir nuevos deseos no derivados y destruir viejos deseos. Pues, como él dice, un agente «puede pensar que tiene una razón para promover algún desarrollo porque no ha ejercitado su imaginación lo suficiente para ver qué pasaría si lo hiciera», y, más «positivamente, la imaginación puede crear nuevas posibilidades y nuevos deseos». Cuando la imaginación crea y destruye deseos de estas maneras, WILLIAMS nos dice que consideramos sus operaciones como dictadas por la razón.

WILLIAMS está en lo cierto, creo, al decir que la deliberación puede tanto producir nuevos deseos no derivados como destruir viejos deseos no derivados. Pero yerra al sostener que la única manera, o quizás la más importante, en que esto ocurre es vía el ejercicio de la imaginación. Pues de lejos el modo más importante en que creamos nuevos deseos y destruimos viejos deseos no derivados cuando deliberamos es tratando de descubrir si nuestros deseos están *sistemáticamente justificados*. Y si esto es correcto, entonces, eso requiere a su vez una cualificación significativa de la afirmación de WILLIAMS según la cual la razón dicta la operación de la imaginación.

¿Qué quiero decir cuando digo que a veces deliberamos al tratar de descubrir si nuestros deseos están sistemáticamente justificados? Quiero decir que podemos intentar decidir si algún deseo no derivado particular que tenemos o podríamos tener es un deseo de hacer algo que a su vez

es deseable no derivadamente. Y lo hacemos de un modo característico: esto es, tratando de integrar el objeto de ese deseo en un perfil desiderativo y una perspectiva evaluativa más coherentes y unificados. RAWLS describe los elementos básicos de este procedimiento de justificación sistemática en su discusión de cómo intentamos encontrar un «equilibrio reflexivo» entre nuestras creencias evaluativas generales y específicas (RAWLS, 1951; DANIELS, 1979). Me circunscribiré a decir algo acerca del modo en que lograr el equilibrio reflexivo también puede ser un objetivo en la formación de deseos no derivados.

Supongamos que tomamos todo un conjunto de deseos que tenemos por cosas generales y específicas. Deseos que no son de hecho derivados de ningún deseo que tenemos de algo más general. Podemos preguntarnos si no podríamos llegar a un conjunto de deseos más sistemáticamente justificable a través de agregar a este conjunto de deseos generales y específicos otro deseo general, o todavía un deseo más general, un deseo que, a su vez, justifique y explique los deseos más específicos que tenemos. Y la respuesta podría ser que sí lo podríamos conseguir. Pues en la medida en que el nuevo conjunto de deseos —el conjunto que imaginamos tener si agregamos un deseo más general a los deseos más específicos que de hecho tenemos— exhibe más, digamos, unidad, podemos decir que el nuevo conjunto imaginario de deseos es más racionalmente preferible al anterior. Pues podríamos considerar apropiadamente la unidad de un conjunto de deseos como una virtud. Una virtud que a su vez hace mejor a la racionalidad del conjunto como un todo. Pues exhibir unidad es parcialmente constitutivo de tener un conjunto de deseos sistemáticamente justificado, y por lo tanto racionalmente preferible, tal como exhibir unidad es parcialmente constitutivo de tener un conjunto de creencias sistemáticamente justificado, y por lo tanto racionalmente preferible.

La idea aquí es directamente análoga a lo que RAWLS tiene que decir sobre las condiciones bajo las cuales podríamos llegar a pensar que debemos adquirir una nueva creencia en un principio más general dado nuestro *stock* de creencias evaluativas específicas. Podríamos descubrir que nuestros juicios de valor específicos serían justificados y explicados de un modo más satisfactorio si los viéramos como abarcados por un principio más general. El conjunto de creencias imaginarias al que llegamos agregando la creencia en el principio más general puede exhibir más unidad que nuestro *stock* actual de creencias, tal como nuestro conjunto imaginario de deseos puede exhibir más unidad que nuestro conjunto actual de deseos.

Si decidimos que nuestros deseos más específicos están mejor justificados, y por lo tanto están mejor explicados, de este modo, entonces adviértase que eso por sí mismo causa que tengamos un nuevo deseo, no derivado, de

aquella cosa más general. Y si lo hace, parece completamente correcto y apropiado suponer que hemos arribado a este nuevo deseo por un método racional. De hecho, la adquisición del deseo nuevo más general parecerá requerida racionalmente exactamente del mismo modo en que la adquisición de una nueva creencia según la cual el objeto del deseo es deseable será considerada requerida racionalmente. De hecho, si el análisis de deseabilidad que ofrecí aquí está bien orientado, la adquisición de una nueva creencia evaluativa será la contraparte cognitiva de la adquisición de un nuevo deseo. Pues —si el análisis es correcto— una creencia evaluativa es simplemente una creencia acerca de lo que sería deseado si fuésemos completamente racionales, y el nuevo deseo es adquirido precisamente porque se lo cree requerido para que seamos racionales.

Más aun, adviértase que si esto es aceptado entonces no solamente podemos explicar cómo podríamos adquirir nuevos deseos como resultado de tal reflexión, sino que también podemos explicar cómo podemos perder viejos deseos también. Pues, dado el objetivo de tener un conjunto de deseos sistemáticamente justificable, bien puede suceder que, a medida que procede el intento por la justificación sistemática, ciertos deseos que de otro modo parecían inobjetables deben ser abandonados. Quizás porque no encontramos ningún modo de integrar esos deseos en un conjunto como un todo, empezarán a aparecer *ad hoc* y por lo tanto injustificables. Nuestra creencia de que tales deseos son *ad hoc* puede entonces causar que perdamos esos deseos. Y, si así ocurre, entonces parecerá sensato describir esto como una pérdida que está dictada por la razón; otra vez, de un modo directamente análogo a la pérdida de una creencia injustificable en tanto *ad hoc*.

A medida que este proceso de justificación sistemática continúa podemos entonces imaginar cambios enteros en nuestro perfil desiderativo. El razonamiento sistemático crea nuevos deseos y destruye viejos. Dado que cada uno de esos cambios parece racionalmente requerido, el nuevo perfil desiderativo no sólo parecerá diferente del antiguo, sino mejor. Más racional. De hecho, parecerá mejor y más racional exactamente del mismo modo, y por las mismas razones, que nuestras nuevas creencias evaluativas correspondientes serán vistas mejor y más racionales que las viejas.

Esto, entonces, es lo que quiero decir al afirmar que podemos crear nuevos deseos y destruir viejos deseos no derivados intentando lograr un conjunto de deseos sistemáticamente justificable. Si lo que he dicho sobre el razonamiento sistemático parece correcto, entonces tiene que ser claro que la afirmación de WILLIAMS según la cual la razón está del lado de la operación de la imaginación requiere una cualificación significativa. Pues si bien la imaginación puede producir de hecho nuevos deseos y destruir viejos *vía* la presentación vívida de los hechos, no está garantizado que sus

operaciones produzcan y destruyan deseos que ellos mismos serían dictados por un intento de justificación sistemática del tipo que se acaba de describir. De hecho, más bien lo contrario es verdadero, pues la imaginación es proclive a todo tipo de influencias distorsivas. Influencias que el razonamiento sistemático tiene la función de eludir. Por ejemplo, al imaginar vívidamente cómo sería matar a alguien, podría considerarme a mí mismo completamente reacio al prospecto de matar a alguien independientemente de su resultado. Pero, a pesar de eso, bien podría encontrar que el deseo de matar a alguien, dados ciertos resultados, es un elemento en un conjunto de deseos sistemáticamente justificable. [Para críticas similares vea Mark JOHNSTON (1989) comentando a David LEWIS (1989) sobre el rol de la imaginación en la deliberación].

El rol jugado por los intentos de justificación sistemática es entonces lo que es crucialmente requerido para una comprensión de cómo la deliberación crea deseos nuevos y destruye deseos viejos no derivados, no el rol jugado por la imaginación. De aquí en más daré por sentado entonces que la condición iii) de WILLIAMS, la condición de la deliberación correcta, debe entenderse de ese modo.

Quinta pregunta: «¿Puede el análisis hacerse más completamente reductivo y explícito?».

El análisis de las razones normativas que he ofrecido hasta aquí —el análisis en términos de hechos acerca de lo que desearíamos en condiciones de racionalidad plena— es evidentemente un análisis de estilo sumario no reductivo de nuestro concepto de una razón normativa. Es un análisis de estilo sumario porque la idea de ser completamente racional es ella misma una idea sumaria; de hecho, acabamos de ver qué es lo que resume. Y es un análisis no reductivo porque al elaborar nuestra idea de qué es ser completamente racional, hemos tenido que usar conceptos normativos. Esto era evidente en nuestra descripción del razonamiento sistemático; pero al dar tal descripción dijimos que apuntamos a encontrar un perfil desiderativo más «unificado», y esa noción es ella misma, por supuesto, una noción normativa.

La pregunta, entonces, es si podemos transformar el análisis en un análisis completamente explícito y reductivo; lo que antes llamé un análisis «en red» (capítulo II). Pero antes de intentar responder a esta pregunta me parece que haríamos bien en recordar que si es posible o no transformar el análisis que se ofrece en un análisis en red, eso es algo simplemente no requerido —al menos no lo es en la medida en que nuestro objetivo es simplemente brindar un análisis—. Pues simplemente no hay un requerimiento de que los análisis sean completamente explícitos y reductivos. Sin embargo, habiendo comprendido eso, consideremos la pregunta en sus propios términos.

Para construir un análisis en red de nuestro concepto de una razón normativa deberíamos ser capaces de hacer tres cosas. Primero, deberíamos ser capaces de describir todas las trivialidades que son constitutivas de nuestra idea de ser completamente racional como una larga conjunción. Ninguna de ellas puede ser una descripción de estilo sumario, como lo es la condición iii) de WILLIAMS, su condición de «la deliberación correcta». Segundo, deberíamos ser capaces de quitar toda mención de términos normativos de esas trivialidades, reemplazándolas con variables ligadas. Y entonces, en tercer lugar, deberíamos ser capaces de usar esta descripción abstracta de un conjunto de relaciones para dar definiciones simultáneas de todos nuestros conceptos normativos, un conjunto de definiciones de nuestros conceptos normativos en términos de las relaciones que los varios rasgos normativos tienen entre sí y con el mundo.

Como vimos antes, sin embargo, cuando un conjunto de conceptos está *en gran medida* interdefinido —esto es, cuando el concepto es en parte aprendido a través de la presentación de paradigmas y por lo tanto hay muy poco afuera del círculo de conceptos que están siendo analizados que jueguen un rol en la definición explícita de ninguno de ellos, como, por ejemplo, vimos en el caso de nuestros conceptos de color—, entonces, cuando intentamos dar un análisis en red, aparece el problema de la permutación. Pues si nos concentramos en el caso en cuestión, emerge la posibilidad de que cuando removamos *todos* nuestros conceptos normativos del enunciado de las trivialidades y los reemplacemos con variables ligadas, no tendremos información *suficiente* para encontrar correctamente la extensión de los conceptos normativos que queremos analizar. Podríamos tener una descripción abstracta de un conjunto de relaciones, un conjunto de relaciones que podrían ser ejemplificadas igualmente por, digamos, un conjunto de creencias y deseos razonables y un conjunto de creencias y deseos no razonables. En esta medida, entonces, podemos ser incapaces de usar esta descripción abstracta de un conjunto de relaciones para fijar correctamente la extensión de lo razonable.

Ahora permítaseme confesar que no sé con seguridad si los análisis en red de nuestros conceptos normativos son vulnerables a un problema de la permutación. Pues establecerlo requeriría una demostración de un tipo u otro, y no sólo no ha existido nunca tal demostración, dado que sería una tarea superhumana escribir un enunciado explícito, de estilo no normativo, de las trivialidades que capturan nuestra idea de qué es ser racional, sino que hay buenas razones para pensar que tampoco se encontrará tal demostración. Sin embargo, en virtud de las mismas razones que di cuando consideramos la pregunta paralela en el caso de nuestros conceptos morales, si se me obligara a especular, mi propia evaluación sería que los análisis en red de nuestros conceptos normativos son de hecho vulnerables a un problema

de la permutación. Mis razones son dos. De nuevo, son las mismas que mis razones en el caso moral.

Primero, lo que la discusión de los conceptos de color muestra es que los problemas de permutación emergen cuando un conjunto de conceptos, adquiridos *inter alia* a través de la presentación de paradigmas, está por tanto en gran medida interdefinido. Los problemas de la permutación emergen cuando hay muy pocos conceptos fuera del círculo de conceptos a ser definidos que juegan un rol significativo en las trivialidades que usamos al enunciar una definición explícita de aquellos que están dentro del círculo. Y, por supuesto, esto es justamente lo que encontramos con nuestros conceptos normativos. Están en *gran medida* interdefinidos. Muy poco afuera de la esfera de lo normativo es requerido para definir lo normativo. Y de nuevo, como sucede con nuestros conceptos de color, esto es así porque aprendemos nuestros conceptos normativos al ser presentados con paradigmas —paradigmas de buenos argumentos, de qué significa que una proposición brinde apoyo a otra, etc.— a partir de los cuales aprendemos a generalizar.

Segundo, me parece que tenemos otras razones inductivas para pensar que los análisis en red de nuestros conceptos normativos son vulnerables a un problema de la permutación también. Pues es un hecho significativo sobre la historia de la filosofía que los análisis de conceptos normativos en términos no normativos han sido fracasos espectaculares. Parece que cualquier análisis de ese tipo es vulnerable a la objeción «¿Y qué?» (JOHNSTON, 1989; GIBBARD, 1990: capítulo I). Lo que hace falta para explicar este hecho notable es alguna razón de principio de por qué los conceptos normativos eluden el análisis no normativo. Y la conjetura obvia es que los análisis en red de nuestros conceptos normativos son vulnerables a un problema de la permutación. Pues éste es justamente el tipo de razón de principio que es requerida.

Concluyo, entonces, que el análisis que he ofrecido de nuestra razón normativa es, y permanecerá siempre, un análisis no reductivo de estilo resumen. Pero no es malo por ello. Pues, como hemos visto, simplemente no es requerido que nuestros análisis sean reductivos y explícitos.

Sexta pregunta. «¿El análisis brinda una concepción relativa o no relativa de las razones normativas?».

Dije que quería considerar una consecuencia que WILLIAMS extrae de su propio análisis de las razones normativas. La consecuencia que él extrae es que su análisis respalda una concepción *relativa* de las razones. Él expone la cuestión de este modo.

[L]a verdad de la oración [...] [«A tiene una razón para hacer ϕ »] [...] implica, muy imprecisamente, que A tiene algún motivo que será servido o

promovido por su hacer φ , y si esto no es así entonces la oración es falsa: hay una condición relativa a los objetivos del agente, y si ella no es satisfecha no es verdadero decir [...] que él tiene una razón para hacer φ (1980: 101).

Y después WILLIAMS dice:

Básicamente, y por definición, [...] [un análisis de las razones] [...] debe desplegar una relatividad de [...] [un] [...] enunciado de razón al *conjunto motivacional subjetivo* del agente [...] (1980: 102).

¿Pero por qué dice WILLIAMS esto? Pues, como hemos visto, aun de acuerdo con su análisis la afirmación de que un agente tiene una razón normativa para hacer φ no es una afirmación sobre sus deseos *reales*, sino más bien una afirmación sobre sus deseos *hipotéticos*. La verdad de la oración «A tiene una razón para hacer φ » no implica por lo tanto, ni siquiera «muy imprecisamente», que A *tenga* algún motivo que será servido o promovido por su hacer φ , lo que implica más bien que *tendría* algún deseo de ese tipo si fuera completamente racional.

WILLIAMS podría conceder esto. Pero, podría decir él, eso no muestra que esté equivocado cuando afirma que un análisis de las razones debe desplegar una relatividad de las razones del agente a sus deseos reales, sólo muestra que el tipo de relatividad en cuestión requiere una formulación más cuidadosa. El punto crucial es que los deseos que un agente tendría si fuese completamente racional son ellos mismos simplemente funciones de sus deseos reales, donde las funciones relevantes son aquellas descritas en las condiciones i) a iii). Las razones de un agente son así relativas a sus deseos reales, él podría decir, porque no podemos esperar que, aun bajo condiciones de racionalidad plena, todos los agentes converjan en los mismos deseos acerca de lo que debe hacerse en las distintas circunstancias que podrían enfrentar. Aun si es racional para cada uno de nosotros cambiar nuestros deseos reales intentando llegar a un conjunto de deseos que pueda ser justificado sistemáticamente —en la manera capturada por las condiciones i) a iii)— tales cambios siempre se quedarán cortos respecto de tener los mismos deseos que nuestros colegas. Siempre reflejarán el hecho antecedente de que tenemos los deseos reales que tenemos.

Como yo lo veo, ésta es la mejor interpretación de la afirmación de WILLIAMS de que nuestras razones son todas relativas. Esto explica por qué está bastante en lo cierto al insistir en que está defendiendo una concepción «humeana» de las razones normativas (1980: 102). Pues su concepción de las razones, como las de HUME, predica un escepticismo sobre el alcance del cambio razonado en nuestros deseos (KORSGAARD, 1986). Predica negar que, a través de un proceso de deliberación racional —a través de intentar dar

una justificación sistemática de nuestros deseos, por ejemplo — podríamos alguna vez llegar a descubrir razones que todos compartimos. Pues lo que tenemos razón para hacer viene dado por el contenido de los deseos que tendríamos si fuésemos completamente racionales, y esos deseos difieren en su contenido de un agente a otro. La visión humeana de WILLIAMS es por lo tanto opuesta a la visión antihumeana o kantiana según la cual en condiciones de racionalidad plena todos razonaríamos del mismo modo hasta llegar a las mismas conclusiones sobre qué debe hacerse; es opuesta a la visión según la cual, *vía* un proceso de justificación sistemática de nuestros deseos, podríamos lograr que converjamos en los deseos que tenemos.

La pregunta a responder es por lo tanto si WILLIAMS está en lo cierto al afirmar que nuestro concepto de una razón normativa presupone ese escepticismo sobre el alcance del cambio razonado en nuestros deseos. ¿Presupone nuestro concepto de razón normativa que habrá, o alternativamente que no habrá, una convergencia en los deseos que tendríamos en condiciones de racionalidad plena? Si presupone que no habrá tal convergencia entonces nuestro concepto de razón normativa es relativo. Si presupone en cambio que habrá tal convergencia entonces nuestro concepto de una razón normativa es, por contraste, no relativo.

En términos de la distinción introducida antes, adviértase que estamos formulando una pregunta *conceptual*, no una pregunta *sustantiva* (capítulo III). Estamos preguntando qué queremos decir cuando decimos que el agente es completamente racional; si parte de lo que queremos decir con «racional» es que las personas plenamente racionales convergen en sus deseos, o si esto no es parte de lo que queremos decir con «racional». Sin embargo, aunque estamos formulando una pregunta conceptual, adviértase que no por ello estamos presuponiendo ninguna postura sustantiva. Aun si nuestro concepto de una razón normativa es él mismo no relativo —aun si nuestro concepto presupone optimistamente que todos convergiríamos en los mismos deseos en condiciones de racionalidad plena— el mundo podría desilusionarnos. Diferencias atrincheradas y racionalmente inexplicables aparentemente en lo que deseamos podrían hacer imposible creer, sustantivamente, que hay tales razones normativas no relativas. Volveré a esta idea en el capítulo siguiente.

Confrontemos, entonces, la cuestión conceptual directamente. ¿Nuestro concepto de razón normativa es relativo o no relativo? La relatividad de una afirmación debería manifestarse en el modo en que hablamos. Considérese, por ejemplo, la afirmación esquemática «es deseable que p en las circunstancias C». En la concepción no relativa de las razones normativas —al menos si dejamos a un lado algunas complicaciones con las que lidiaré en un momento— esta afirmación tiene una condición de verdad directa: es

deseable que p en C justo en el caso de que todos deseáramos que p en C si fuésemos plenamente racionales.

Hay, por lo tanto, un sentido en que podemos hablar de justificación racional o deseabilidad *simpliciter*. Cuando usted y yo hablamos de las razones que hay para actuar, estamos por lo tanto hablando sobre lo mismo. Estamos hablando sobre razones y punto.

En la concepción relativa, sin embargo, el asunto es bastante diferente. Pues para dar la condición de verdad de la afirmación esquemática «es deseable que p en C» debemos saber primero desde qué perspectiva ha de evaluarse la verdad de la afirmación. Pues mientras que «es deseable que p en C» en tanto evaluada desde la perspectiva de A es verdadera si y sólo si A desearía que p en C si A fuese plenamente racional; «es deseable que p en C» en tanto evaluada desde la perspectiva de B es verdadera si y sólo si B desearía que p en C si B fuese plenamente racional, y así sucesivamente. No hay por lo tanto nada como la deseabilidad o justificación racional *simpliciter*, pero sólo deseabilidad_A, deseabilidad_B, [...] La-justificabilidad-racional-desde-la-perspectiva-de-A, la-justificabilidad-racional-desde-la-perspectiva-de-B..., etc. Si le digo «hay una razón para hacer ϕ » y usted lo niega, estamos por lo tanto potencialmente hablando sobre cosas diferentes: las razones_{yo} y las razones_{usted}. La pregunta a formular es entonces si el modo en que hablamos de las razones para la acción y la justificación racional refleja una concepción relativa o no relativa de las condiciones de verdad.

Una razón para pensar que refleja una concepción no relativa viene del contexto más amplio en el cual la pregunta es formulada. Pues es importante recordar que tenemos un amplio rango de conceptos normativos: verdad, significado, apoyo, implicación, deseabilidad, etc. Entre ellos estos conceptos nos permiten formularnos todo tipo de preguntas normativas, preguntas acerca de lo que deberíamos y no deberíamos creer, decir y hacer. ¿Pero cómo es que muchos de estos conceptos normativos puedan ser plausiblemente considerados como dando lugar a que las afirmaciones tengan condiciones de verdad relativizadas? Como yo lo entiendo, ninguna lo hace.

Considérese nuestro concepto de apoyo, a modo de ejemplo. Parece bastante implausible suponer que la verdad de afirmaciones sobre qué proposiciones brindan apoyo a otras es implícitamente relativa al individuo. Cuando A dice « p brinda apoyo a q » y B dice « p no brinda apoyo a q » están hablando potencialmente sobre cosas bien diferentes: que A está hablando sobre que brinda apoyo_A a q y B está hablando sobre que brinda apoyo_B a q , por ejemplo. Pues si ese fuera el caso entonces deberíamos darnos cuenta de que a veces es posible disolver desacuerdos aparentes al descubrir que ambas partes están hablando sinceramente. Debería ser permisible para B decir «A dijo que “ p brinda apoyo a q ” y lo que ella dijo es verdadero, pero p no brinda apoyo a q ». Sin embargo, es un

rasgo notable de nuestro discurso sobre qué proposiciones brindan apoyo a otras que *nunca* disolvemos los desacuerdos aparentes de este modo. Las proposiciones tienen fuerza normativa *simpliciter*, no simplemente fuerza-normativa-relativa-a-este-individuo o relativa-a-aquel-individuo. Cuando un individuo dice « p brinda apoyo a q » y el otro dice « p no brinda apoyo a q » entonces expresan su desacuerdo acerca de si p brinda apoyo a q en un sentido *no relativo*.

Si nuestro concepto de deseabilidad fuese implícitamente relativizado, entonces parece que ello marcaría una diferencia significativa entre este concepto y nuestros otros conceptos normativos. Deberíamos esperar encontrar que con afirmaciones acerca de lo que es deseable, a diferencia de lo que sucede con afirmaciones acerca de qué proposiciones brindan apoyo a cuáles, *somos* capaces de disolver los desacuerdos aparentes del modo descrito ¿Pero es esto así?

Podría pensarse que sí. Después de todo, ¿no hay acaso todo un conjunto de casos familiares en que decimos cosas como «eso puede ser una razón para usted pero no para mí», «quizás sea deseable para usted, pero no es deseable para mí», y cosas por el estilo? Pero aunque tales casos de hecho existen, es importante advertir que el tipo de relatividad que señalamos cuando decimos tales cosas es bien diferente del tipo recién descrito. Bien diferente del tipo de relatividad que WILLIAMS tiene en mente. Pues, en los casos familiares, «eso puede ser una razón para usted pero no para mí» señala el hecho de que hay una relatividad incluida en las *consideraciones* que usamos para justificar racionalmente nuestras elecciones. No señala el hecho de que nuestro concepto de justificación racional sea él mismo relativo al individuo; que no haya tal cosa como consideraciones, relativas o no, que justifican racionalmente nuestras elecciones, sino consideraciones que justifican-racionalmente-relativas-a-esta-persona o que justifican-racionalmente-relativas-a-aquella-persona. Aquí, entonces, llegamos a las complicaciones que pusimos a un lado antes.

Algunas veces lo que tenemos en mente cuando decimos «eso puede ser una razón para usted pero no para mí» es que las consideraciones que justifican racionalmente nuestras elecciones son, para usar los términos de PARFIT, *relativas-al-agente* más que *neutrales-en-relación-con el-agente* (PARFIT, 1984). Supongamos que usted está en una playa. Dos personas se están ahogando a su izquierda y una se está ahogando a su derecha. Puede o bien nadar hacia la izquierda y salvar a las dos personas, en cuyo caso la persona de la derecha se ahogará, o puede nadar hacia la derecha y salvar una, en cuyo caso las dos personas de la izquierda se ahogarán. Usted decide nadar hacia la derecha y salvar a una sola persona y justifica su acción diciendo «la persona que está a la derecha es mi hijo, mientras que los otros dos a la izquierda son perfectos extraños para mí».

En un sentido, por supuesto, yo bien puedo decir «ésta puede ser una razón para usted pero no lo es para mí». Pues si las tres personas que se ahogan son todos perfectos extraños para mí, entonces, si yo hubiera estado parado en la playa en vez de usted, no hubiera sido capaz de justificar mi elección de nadar hacia la derecha y salvar sólo a una. Pero en otro sentido parece que lo que es una razón para usted puede de hecho ser una razón para mí. Pues si yo hubiera estado parado en la playa en vez de usted y la persona a mi derecha hubiese sido mi hijo, entonces seguramente habría sido capaz de justificar mi elección de nadar hacia la derecha y salvar solo a una persona diciendo «el de la derecha es mi hijo». De hecho, si pensamos que un padre que fracasa en salvar a su hijo en tales circunstancias, fracasa en actuar en función de una razón disponible para él —como me parece que hacemos— entonces estamos de hechos forzados a decir esto. Forzados a asumir la concepción no relativa de las razones normativas.

Lo que este tipo de ejemplo muestra es entonces que, aun si las razones son no relativas en el sentido crucial en juego aquí, entre las consideraciones que pueden justificar racionalmente nuestras elecciones hay tanto consideraciones que pueden ser apropiadamente formuladas *de dicto* como consideraciones que pueden ser apropiadamente formuladas *de se* (véase también LEWIS, 1989). Esto es, hay tanto razones *de dicto* como *de se*. Ambos podemos expresar el contenido de la razón *de dicto* relevante en este caso usando las palabras «hay una razón para salvar a la gente en general» y cada uno puede expresar el contenido de la razón *de se* usando las palabras «hay una razón para salvar a mi hijo en particular». En estos términos podemos decir que lo que es una razón para usted, en este caso, no es una razón para mí en el sentido de que, si hubiese sido yo quien estaba parado en la playa en vez de usted, y si las mismas personas se hubieran estado ahogando, entonces la única consideración que habría sido relevante para mi elección es la razón *de dicto*. La razón *de se* no habría sido relevante para mi elección porque la gente que de hecho se estaba ahogando eran perfectos extraños para mí. Pero en otro sentido lo que es una razón para usted es de hecho una razón para mí. Pues si yo hubiera estado parado en la playa y la persona a la derecha hubiese sido mi hijo, tal como la persona a la derecha es su hijo, entonces tanto la razón *de dicto* como la *de se* habrían sido relevantes para mi elección exactamente del mismo modo en que ambas son relevantes para la suya.

Dije que este tipo de relatividad es diferente del tipo de relatividad que WILLIAMS tiene en mente, y ahora debería ser claro por qué ello es así. Pues, en términos del análisis, aun si algunas consideraciones que justifican racionalmente nuestras elecciones son relativas porque son *de se*, la existencia de razones *de se* todavía puede requerir una convergencia en los deseos que todos tendríamos si fuésemos plenamente racionales. Esto es, la existencia

de razones con contenidos *de se* puede todavía requerir que, en condiciones de racionalidad plena, cada uno de nosotros tenga deseos cuyos contenidos expresaríamos usando palabras como «ayudar a mis hijos», «promover mi bienestar», y cosas por el estilo. La mera existencia de razones *de se* es por lo tanto bastante diferente de la relatividad que WILLIAMS tiene en mente. Pues su afirmación es que las razones son relativas en el sentido de que no requieren tal convergencia; el hecho de que mi acto ayude a mi hijo puede justificar-racionalmente-relativo-a-mí, aun si el hecho de que su acto ayude a su hijo no justifique-racionalmente-relativo-a-usted.

Hay otro tipo familiar de relatividad en nuestras afirmaciones sobre las razones que también tenemos, un tipo de relatividad que se deriva del hecho de que lo que tenemos razón para hacer es relativo a nuestras circunstancias, donde nuestras circunstancias pueden incluir aspectos de nuestra propia psicología. Supongamos, por ejemplo, que usted y yo diferimos en nuestras preferencias por el vino o la cerveza. Dado que usted prefiere el vino, usted bien podría decirme que hay una razón para ir a la tienda de vinos después del trabajo, pues allí se vende muy buen vino. Pero entonces, dado que yo prefiero la cerveza, yo puedo correctamente replicar «esa puede ser una razón para usted para ir a la tienda de vinos, pero no es una razón para mí».

Ahora bien, esto podría parecer inicialmente similar a la afirmación de que nuestras razones son relativas a nuestros deseos en un sentido similar al que WILLIAMS tiene en mente, pero en realidad no. Pues el punto crucial en este caso es que un rasgo relevante de sus circunstancias es su preferencia por el vino, mientras que un rasgo relevante de mis circunstancias es mi preferencia por la cerveza. Que esto es un rasgo relevante de nuestras circunstancias es manifiesto a partir del hecho de que puedo acordar alegremente con usted en que si yo fuera usted en sus circunstancias —si yo prefiriese el vino a la cerveza— entonces el hecho de que la tienda de la esquina venda muy buen vino constituiría una razón para mí para ir ahí también, tal como constituye una razón para usted.

Este tipo de relatividad es por lo tanto completamente diferente del tipo de relatividad que WILLIAMS tiene en mente también. Pues, en términos del análisis, aun si las preferencias de un agente entran en una especificación de las circunstancias que el agente enfrenta, todavía puede ser el caso de que si el agente está racionalmente justificado al considerar sus propias preferencias, y el modo en que está justificado en tomarlas en cuenta si lo está, depende de si los agentes plenamente racionales convergirían todos en un deseo que hace que las preferencias que él tiene sean relevantes para su elección, y, si lo hacen, el modo en que el deseo en que convergen hace que sus preferencias sean relevantes a su elección (PETTIT y SMITH, 1993b). El hecho de que, al justificar racionalmente nuestras elecciones, a veces nuestras preferencias

sean un rasgo relevante de nuestras circunstancias no respalda entonces la idea de WILLIAMS de que la justificación racional es ella misma una cuestión relativa. Que realmente sólo hay una justificación-racional-relativa-a-esta-persona o una justificación-racional-relativa-a-aquella-persona.

Para encontrar apoyo para el tipo de relatividad que WILLIAMS tiene en mente, necesitamos por lo tanto buscar casos en que sea permisible hacer afirmaciones mucho más radicalmente relativizadas sobre lo que hay una razón para hacer. Pero de hecho, hasta donde puedo ver, no encontramos afirmaciones de ese tipo. Supongamos que alguien me dice que tiene una razón para tomar vacaciones y que yo pienso que no tendría ninguna razón para tomar vacaciones en las circunstancias que ella enfrenta. Suponiendo que hemos tenido en cuenta apropiadamente consideraciones *de se* que podrían ser relevantes para su elección, y suponiendo que hemos tenido apropiadamente en cuenta el modo en que sus preferencias pueden constituir un rasgo relevante de sus circunstancias, me parece que directamente desacuerdo con ella sobre la justificabilidad racional de que ella tome vacaciones en las circunstancias que ella enfrenta, un desacuerdo que yo puedo expresar diciendo «ella piensa que hay una razón para tomar unas vacaciones en sus circunstancias pero no hay tal razón». Si ella cita una consideración en favor de tomar unas vacaciones que yo pienso que no lo justifica, entonces no concluyo que eso puede justificar-relativo-a-ella, aunque no justificar-relativo-a-mí. Yo concluyo que no lo justifica *simpliciter*.

El punto es importante, pues sugiere que cuando hablamos sobre las razones para la acción normalmente nos consideramos hablando sobre un asunto común. Estamos por lo tanto en acuerdo o en desacuerdo entre nosotros acerca de qué constituye una razón y qué no. Ésta es la razón por la que, cuando nos encontramos en desacuerdo —como por ejemplo en el caso de desacuerdo sobre si hay o no una razón para tomar vacaciones en ciertas circunstancias— siempre tenemos la opción de involucrarnos en una discusión en el intento para encontrar quién está en lo cierto y quién está equivocado. Las opiniones de otras personas sobre las razones que hay constituyen entonces desafíos potenciales a mis propias opiniones. Tengo algo que aprender sobre mí mismo y sobre mis propias razones al darme cuenta de cosas sobre los otros y sus razones. Ésa es la razón por la cual los libros y las películas son tan absorbentes. Todo esto es directamente inconsistente con la afirmación de que nuestro concepto de una razón para la acción es en general relativo al individuo; que típicamente significa razón_{para mí} cuando yo lo digo, razón_{para ti} cuando tú lo dices, razón_{para ella} cuando ella lo dice, etc. Sugiere más bien que nuestro concepto de razón normativa es tozudamente no relativo.

De hecho, me parece que no tenemos más opción que pensarlo así; pues si las razones normativas fueran de hecho relativas, entonces la mera reflexión

sobre ese hecho sería suficiente para socavar su significatividad normativa. Pues en la concepción relativa resulta que, por ejemplo, la deseabilidad^{relativa} a mí de alguna consideración *p* es enteramente dependiente del hecho de que *mis* deseos reales son tales que, si *yo* fuera a involucrarme en un proceso de justificación sistemática de *mis* deseos, quitando a un lado a aquellos que no están justificados y adquiriendo a aquellos que sí lo están, un deseo de que *p* sería uno de los deseos que *yo* terminaría teniendo. Pero cuáles son *mis* deseos reales con los que empiezo, en esta concepción relativa de las razones, es un asunto completamente *arbitrario*, un asunto sin ninguna significación normativa propia. Podría haber tenido cualquier conjunto de deseos viejos al empezar, aun un deseo del que se desprendiera el deseo de que no *p* después de un proceso de justificación sistemática. La deseabilidad^{relativa a mí} del hecho de que *p* entonces termina siendo un hecho completamente arbitrario. Pero la arbitrariedad es precisamente un rasgo de una consideración que tiende a socavar cualquier significación normativa que podría haber parecido tener inicialmente (SMITH, 1989; DARWALL, GIBBARD y RAILTON, 1992).

En la concepción no relativa, por contraste, la reflexión sobre nuestro concepto de deseabilidad no revela tal arbitrariedad. Pues en esa concepción cualquiera puede razonar hacia los mismos deseos si se involucra en un proceso de justificación sistemática de sus deseos. Qué deseos *yo* terminaría teniendo, después de involucrarme en tal proceso, no depende por tanto de modo alguno de cuáles son *mis* deseos reales iniciales. La razón misma determina el contenido de nuestros deseos completamente racionales, no el hecho arbitrario de que tenemos los deseos reales que tenemos. La reflexión sobre el concepto de deseabilidad deja así perfectamente intacta la significación normativa de hechos sobre lo que es deseable e indeseable.

He estado argumentando que la verdad de una afirmación de razón normativa requiere la convergencia en los deseos de los agentes plenamente racionales. Sin embargo adviértase que la convergencia requerida no está al nivel de los deseos sobre cómo cada agente ha de organizar su propia vida en su propio mundo. En sus propios mundos los agentes plenamente racionales se encontrarán en circunstancias bien diferentes entre sí, circunstancias que están condicionadas por sus diferentes encarnaciones, talentos, entornos y afectos en sus mundos respectivos. Sus deseos sobre cómo organizar sus propias vidas en sus propios mundos reflejarán por lo tanto estas diferencias en sus circunstancias. La convergencia requerida está más bien al nivel de sus deseos hipotéticos acerca de qué ha de hacerse en las distintas circunstancias en que podrían encontrarse a sí mismos.

El mero hecho de que una convergencia en los deseos hipotéticos de criaturas plenamente racionales sea requerido para la verdad de afirmaciones de razón normativa no garantiza que tal convergencia vaya a llegar, por

supuesto. Al defender la *concepción* no relativa de las razones normativas no hemos dicho por lo tanto nada para sugerir que, *sustantivamente*, hay tales razones. Pero lo que hemos dicho sí sugiere que, para descubrir si hay razones normativas, y en su caso cuáles son, no tenemos más alternativas que brindar argumentos y ver adónde conducen. La convergencia sustantiva siempre se asume disponible, en la medida en que conversamos y discutimos sobre las razones que tenemos. Pero si este supuesto es verdadero o no, siempre está *sub judice*; es algo que ha de ser descubierto en función del resultado de esas conversaciones y argumentos (compárese PITT sobre seguimiento de reglas 1993: esp. 96-97).

Séptima pregunta. «¿Son las afirmaciones de razón normativa imperativos categóricos o imperativos hipotéticos?».

Foot afirma que los requerimientos de la razón práctica son imperativos hipotéticos. Pues, nos dice ella, las afirmaciones sobre lo que un agente tiene una razón para hacer deben ser retiradas «si encontramos que la relación correcta no se mantiene entre la acción y el fin —que o bien no es un modo de lograr lo que él quiere (o de hacer lo que él quiere) o que no es el medio más elegible entre los distintos posibles—». U otra vez porque nos dice que el agente puede rechazar una afirmación de razón normativa «mostrando que su acción no es subsidiaria de sus intereses o deseos» (Foot, 1972: 159). Nada de esto sería posible si las afirmaciones de razón normativa fueran imperativos categóricos. Pues entonces la verdad de la afirmación de razón sería una función de las circunstancias en que el agente se encuentra independientemente de si él tiene un deseo o interés en lo que se supone que tiene una razón para hacer.

Pero, como ahora debería ser claro, la idea de que las afirmaciones de razón son imperativos hipotéticos —al menos tal como Foot caracteriza el imperativo hipotético— simplemente se da de bruces con lugares comunes sobre los modos en que las razones normativas y las razones motivacionales pueden divergir. Pues es un lugar común que, por ejemplo, los depresivos tienen razones para hacer todo tipo de cosas. Pero éstas no son razones para hacer cosas que ellos desean o en las que tengan interés, pues el efecto de su depresión ha sido precisamente destruir cualquier deseo o interés que hayan tenido en algo, aun las cosas respecto de las cuales ellos creen que tienen una razón para hacer. Las afirmaciones sobre las razones que los depresivos tienen no deben por lo tanto ser retiradas simplemente porque los depresivos carezcan de los deseos relevantes.

En el mejor de los casos, entonces, para que un agente rechace la afirmación de que tiene una razón normativa para actuar de cierto modo, lo que tiene que mostrar es que actuar de ese modo no es subsidiario de sus deseos *hipotéticos*: los deseos que tendría, si fuera plenamente racional, acerca de

qué ha de hacer en las circunstancias que enfrenta. Sin embargo, tal como hemos visto, dado que nuestro concepto de razón normativa no es relativo, al evaluar la verdad de tales afirmaciones presuponemos que agentes plenamente racionales tendrían todos los mismos deseos acerca de lo que ha de ser hecho y deseado en distintas circunstancias. En esta teoría, las afirmaciones de razón normativa son por lo tanto imperativos categóricos, pues los agentes que enfrentan las mismas circunstancias todos tienen las mismas razones.

Debemos por lo tanto rechazar la afirmación de Foot según la cual las afirmaciones de razón son imperativos hipotéticos. El carácter no relativo de las razones normativas implica que las afirmaciones de razón normativa son imperativos categóricos después de todo.

Octava pregunta. «¿Cuáles son las consecuencias epistemológicas del hecho de que las afirmaciones de razón normativa sean imperativos categóricos?».

Considérese de nuevo la discusión anterior sobre la deliberación. Vimos entonces que uno de los modos en que podemos decidir lo que tenemos razón normativa para hacer —esto es, lo que desearíamos que hiciéramos en condiciones de racionalidad plena— es intentando encontrar un conjunto de deseos que es sistemáticamente justificable. Pues tal conjunto de deseos será nuestra mejor evaluación de los deseos que tendríamos en condiciones de racionalidad plena. Sin embargo, ahora que hemos visto que la verdad de una afirmación de razón normativa presupone que agentes plenamente racionales tendrían todos los mismos deseos acerca de qué ha de hacerse en distintas circunstancias que enfrenten, esta tarea de encontrar una justificación sistemática de nuestros deseos empieza a tomar una dimensión distintivamente social. Permítaseme explicar por qué.

Dado que las opiniones de otras personas acerca de qué es y no es deseable, no son meramente las mismas o diferentes que las propias, se sigue que ellas potencialmente confirman o refutan nuestras propias opiniones —dependiendo, por supuesto, de la calidad de los argumentos que pueden ofrecerse en su apoyo—. Por lo tanto no es sólo importante para ellos, sino también importante para nosotros, cuáles de sus deseos ellos consideran apropiados como puntos de partida para este proceso de justificación sistemática descrito antes, no sólo porque dónde ellos terminarán será una función de dónde empiezan, sino porque sus juicios sobre el carácter apropiado de los deseos iniciales están ellos mismos potencialmente en conflicto con nuestros propios juicios sobre los deseos a partir de los cuales sería apropiado empezar. Pues, si hay hechos no relativos acerca de qué es deseable, entonces los deseos a partir de los cuales es apropiado para cualquiera empezar a involucrarse en este razonamiento sistemático en un intento por encontrar

esos hechos sólo pueden ser aquellos deseos a partir de los cuales somos capaces de construir una convergencia sustantiva en nuestros deseos.

De modo no sorprendente, entonces, un argumento sobre lo que es y no es deseable se convierte en un argumento como el de qué apoya a qué. La epistemología del valor, como la epistemología de qué apoya a qué, requiere al individuo que se considere a sí mismo como uno entre un grupo de individuos que están tratando de encontrar una respuesta a un conjunto común de preguntas, preguntas cuya formulación no requiere una referencia a uno de ellos en particular. Ella debe admitir que, todo lo demás igual, nadie está mejor posicionado que ningún otro para responder a estas preguntas simplemente en virtud de ser la persona que es: ninguno es infalible sobre tales cuestiones y ninguno es incapaz de tener una opinión que merezca la pena escuchar. Todo lo demás igual, el individuo debe por lo tanto tener un sentido apropiado de humildad cuando se encuentre en desacuerdo con el grupo. Debe admitir que puede adoptar una postura racional en contra del grupo sólo cuando puede construir una historia plausible acerca de por qué su propia opinión es más creíble que la opinión de sus colegas. Y, en función de sus posiciones plausibles, eso no será en general una tarea fácil.

RAWLS estaba él mismo al tanto de todas estas complicaciones. Como yo lo veo, ésa es la razón por la que formuló el equilibrio reflexivo en términos sociales y no individualistas. Por qué, por ejemplo, él insiste en que los juicios apropiados con los cuales iniciar nuestros intentos de justificación sistemática son nuestros «juicios considerados», que él define *inter alia* como sigue:

He requerido que el juicio sea estable, esto es, que haya evidencia de que en otros momentos y en otros lugares jueces competentes han llegado al mismo juicio en casos similares, entendiendo por casos similares aquellos en que los hechos relevantes y los intereses y concurrencias son similares. La similitud tiene que tener lugar, en general, sobre la clase de jueces competentes y sobre sus juicios en diferentes tiempos (RAWLS, 1951: 182).

Al decidir con qué deseos empezar en un intento por encontrar una justificación sistemática de nuestros deseos, entonces, no tenemos más opción que buscar deseos que son de modo similar ampliamente compartidos. No tenemos más opción dados dos supuestos cruciales. Primero, el objetivo es encontrar un conjunto único de deseos que todas las criaturas racionales reconocerían como sistemáticamente justificables, y segundo, que ninguno de nosotros tiene ningún talento epistémico especial que justificaría privilegiar nuestros propios deseos y juicios sobre los deseos y juicios de los otros por adelantado que nos justifique por adelantado en descartar sus deseos y opiniones en contrario como si no tuvieran ninguna significación epistémica para nosotros.

10. EL PROBLEMA RESUELTO

Volvamos al argumento principal. Dije que un análisis de la deseabilidad en términos de qué desearíamos si fuésemos plenamente racionales nos permite dar sentido a C2: la afirmación de que si creo que tengo una razón normativa para hacer φ entonces debería racionalmente desear hacer φ . Ahora estamos en una posición para explicar por qué ello es así.

C2 nos dice que si creemos que tenemos una razón normativa para hacer φ entonces deberíamos racionalmente desear φ . De acuerdo con el análisis, la creencia de que tenemos una razón normativa para hacer φ , o que es deseable hacer φ , puede ser representada como la creencia de qué desearíamos hacer φ si fuésemos plenamente racionales. Pero ahora, supóngase que creemos que desearíamos hacer φ si fuésemos plenamente racionales y fracasáramos en desear hacer φ . ¿Somos irracionales? Ciertamente lo somos. Y en nuestros propios términos y de acuerdo con nuestras propias luces. Pues fracasamos en tener un deseo que creemos que es racional tener. En otras palabras, si creemos que desearíamos hacer φ si fuésemos plenamente racionales entonces deberíamos racionalmente desear φ y esto es justamente C2.

De este modo capturamos la letra de C2, ¿pero podemos capturar su espíritu? ¿Si creemos que desearíamos hacer φ si fuésemos plenamente racionales y sin embargo no deseamos hacer φ , podemos ver por qué deberíamos desprendernos del deseo de no hacer φ y adquirir el deseo de hacer φ en cambio, en vez de, por ejemplo, cambiar nuestra creencia evaluativa? (Acá recordamos el problema de la reducción de valorar a desear desear). Ciertamente podemos.

Nuestro hacer φ es deseable justo en el caso en que desearíamos hacer φ si fuésemos plenamente racionales. Ahora bien, por hipótesis, lo que creemos es que desearíamos hacer φ si fuésemos plenamente racionales. No creemos que desearíamos *no* hacer φ si fuésemos plenamente racionales. Y el mero hecho de que de hecho deseemos no hacer φ no nos da ninguna razón para cambiar esta creencia. No nos da ninguna razón para re-evaluar la verdad de nuestra creencia. Creyendo lo que creemos se sigue por lo tanto que deberíamos racionalmente desprendernos de nuestro deseo de no hacer φ y adquirir en cambio el deseo de hacer φ ².

² ¿Es este argumento consistente con el anterior argumento en contra de JOHNSTON? Sí, lo es. De hecho, ayuda a explicar por qué una teoría como la teoría del autointerés puede autoeliminarse (PARFIT, 1984: 23-24). Supóngase que es racional para mí hacer sólo una cosa: promover mi autointerés a largo plazo. Y supóngase, más aún, que no es racional para mí desear la promoción a largo plazo de mi autointerés; que mi autointerés a largo plazo estaría mejor servido por mi desear actuar por el bien de mi familia y amigos, escribir libros, promover la humanidad, etc., sin tener

Este argumento es, lo admito, muy simple. Como sucede con muchos argumentos simples, su poder real puede por ello ser pasado por alto. Podría ser considerado *demasiado* simple. De modo que permítaseme agregar apoyo adicional a este argumento mostrando que un argumento estructuralmente similar nos permite explicar un fenómeno similar en el caso de la creencia.

Adviértase que el siguiente principio, muy similar a C2, gobierna nuestras creencias:

- C3 Si un agente cree que tiene (la mayor) razón para creer que p entonces debería racionalmente creer que p .

Y adviértase, todavía más, que podemos explicar C3 *vía* un argumento que es estrictamente paralelo al argumento que acabo de dar para explicar C2. Ese argumento descansa sobre una trivialidad sobre las razones para la acción. De manera que considérese una trivialidad sobre nuestras razones para creer. Así como, si tenemos una razón para hacer ϕ podemos decir que ϕ es deseable, donde la deseabilidad está fijada por normas de racionalidad, si tenemos (la mayor) razón para creer que p podemos decir que p es (muy) creíble, donde la credibilidad también está fijada por normas de racionalidad. Pero ahora adviértase que tal como es una trivialidad decir que si hacer ϕ es deseable entonces hacer ϕ es lo que desearíamos si fuésemos plenamente racionales, también es una trivialidad decir que si p es (muy) creíble entonces p es lo que deberíamos creer si fuésemos plenamente racionales. Equipados con estas trivialidades, tenemos lo suficiente para explicar C3.

Suponga que un agente creería que p si fuese plenamente racional y sin embargo fracasa en creer que p . ¿Es irracional? Ciertamente lo es. ¡Y según sus propias luces! Pues fracasa en creer algo que él cree que tiene una (gran) razón para creer. De hecho, este debe seguramente ser un caso paradigmático de irracionalidad. Más aún, adviértase que también podemos explicar por qué alguien que cree que p es (muy) creíble, pero que también encuentra que

ninguna preocupación en modo alguno por mi autointerés de largo plazo. Lo que es deseable que yo haga no es, en este caso, lo que es deseable que yo desee que haga. Pero ahora suponga que me convengo de la teoría del autointerés. Creo que es deseable promover mi auto-interés a largo plazo y que no es deseable que desee promover mi autointerés a largo plazo. En función del argumento recién presentado, tener estas creencias hace que sea racional para mí desear promover mi autointerés a largo plazo y desear no desear promover mi propio autointerés a largo plazo. Dado que la razón de que tengo para promover mi autointerés a largo plazo, algo que sabemos independientemente que racionalmente yo no debería desear, es que yo creo que la teoría del auto-interés, no es sorprendente darse cuenta de que racionalmente yo no debería creer en la teoría del autointerés. La teoría, pues, se autoelimina. Y ya que deseo no desear promover mi autointerés a largo plazo, no sorprende que esté motivado para liberarme de esa creencia. Estoy, en efecto, movido a hacer lo que la teoría me dice que es racional hacer.

cree que no p , racionalmente debería deshacerse de su creencia de que no p y adquirir en cambio la creencia de que p . Pues p es (muy) creíble en caso de que él creería que p si fuese plenamente racional. Y, por hipótesis, eso es lo que él cree. El no cree que creería que no p si fuese plenamente racional. Y el mero hecho de que realmente cree que no p le da una razón para cambiar su creencia. Así, él racionalmente debería deshacerse de su creencia de que no p y adquirir en cambio la creencia de que p . Y eso es justamente C3.

Dada la similitud estructural entre este argumento y el argumento a favor de C2, y dado el éxito del argumento en el caso de la creencia, concluyo que ambos argumentos son exitosos. La trivialidad de que la deseabilidad es una cuestión de lo que desearíamos si fuésemos plenamente racionales es suficiente para mostrar cómo puede ser que nuestras creencias sobre nuestras razones nos requieren racionalmente tener los deseos correspondientes.

La verdad de C2 tiene obvias repercusiones para la naturaleza de la deliberación. Supóngase que un agente que todavía no desea hacer ϕ delibera y, como resultado, llega a creer que tiene una razón normativa de hacer ϕ . Y supóngase adicionalmente que creer eso causa que desee hacer ϕ . Dado C2 se sigue que deberíamos redescubrir esta transición causal entre creencia y deseo en términos normativos. Pues el tener esta creencia causa que él tenga un deseo que es racional que él tenga, dada su creencia. La transición causal entre esa creencia y el deseo se ajusta perfectamente con la transición causal entre, digamos, las creencias de que p y de que p luego q y la creencia de que q . O entre la creencia de que toda la evidencia respalda a q o que es (muy) creíble que q , y la creencia de que q .

Más aún, adviértase que el hecho de que nuestras creencias y deseos tengan tales relaciones normativas entre sí no es inconsistente con la teoría humeana de las razones motivacionales defendida antes. De hecho, toda esta discusión estuvo basada en la teoría humeana. Todas las acciones son de hecho producidas por deseos, tal como el humeano dice. Ninguna acción es producida por creencias solamente o por *creseos*. Pero, si lo que hemos dicho aquí es correcto, algunos de estos deseos son ellos mismos producidos por las creencias del agente sobre las razones que tiene, creencias que él adquiere a través de la deliberación racional³.

³ Una vez que hemos aceptado C2 no tenemos razón para negar que nuestras creencias sobre las razones juegan un rol causal en la producción de nuestros deseos. Sin embargo, nótese que el argumento que he dado para C2 no depende en absoluto de una concepción no relativa de las razones normativas por las que yo abogo. Puede llegar a buen puerto igualmente en la concepción relativa de las razones normativas apoyada por WILLIAMS. Si esto es correcto, entonces se sigue que incluso aquellos que aceptan una concepción relativa, humeana, de las razones, deben admitir que nuestras creencias sobre nuestras razones pueden jugar un papel causal en la producción de nuestros deseos. Esto sugiere que una de las cuestiones que se supone de manera estandarizada divide en humanos y antihumanos a los teóricos de las razones normativas no los divide en ab-

Ahora estamos en una posición para atar todos los cabos sueltos de la discusión. Parecía difícil reconciliar la afirmación de que la deliberación sobre la base de nuestros valores es práctica en su resultado en la medida en que lo es con dos afirmaciones adicionales, la afirmación de que la deliberación normalmente refleja nuestras creencias evaluativas y la afirmación de que nuestras acciones son producidas por nuestros deseos. Sin embargo, hemos visto que estas afirmaciones no están en conflicto. En cambio reflejan un hecho sustantivo sobre los agentes humanos: a saber, que somos criaturas racionales que a veces somos más racionales, a veces menos racionales. La deliberación sobre la base de nuestras creencias evaluativas es práctica en su resultado justo en la medida en que lo es porque *ésta* es precisamente la medida en que somos racionales.

El punto no es que esta respuesta sea de algún modo sorprendente. Siempre fue la única respuesta disponible. Pues si, cuando deliberamos, intentamos decidir lo que tenemos razón para hacer, y en la medida en que somos racionales o bien tendremos ya los deseos correspondientes o bien nuestras creencias acerca de lo que tenemos razón para hacer causarán que tengamos los deseos correspondientes, entonces nada más que el hecho contingente de que somos racionales en la medida en que lo somos podría explicar los ajustes y desajustes resultantes entre nuestras creencias acerca de lo que tenemos razón para hacer y nuestros deseos. Nuestra racionalidad contingente es la única variable. El punto es simplemente que ahora sabemos *por qué* nuestro ser racional juega ese rol. Juega ese rol porque lo que tenemos razón normativa para hacer es una cuestión de lo que desearíamos hacer si fuésemos plenamente racionales.

Se sigue por lo tanto que no hay ningún conflicto en las dos perspectivas de la explicación de la acción descritas al principio: la intencional y la deliberativa. Todas las acciones intencionales son de hecho explicables desde la perspectiva intencional en términos de nuestros deseos y creencias subyacentes. Pero, en la medida en que somos racionales, nuestras acciones también son explicables desde la perspectiva deliberativa, pues nuestros deseos son ellos mismos sensibles a nuestras creencias sobre nuestras razones. Nuestra racionalidad sustantiva explica por lo tanto por qué la conexión entre deliberación y acción no es enteramente contingente y fortuita.

soluto. De acuerdo a Christine KORSGAARD por ejemplo (1986), los antihumanos aseveran, y los humanos niegan, que la razón pueda producir un motivo. Pero, como hemos visto, lo que está en juego no es eso, sino más bien si las razones que producen los motivos son ellas mismas relativas, como suponen los humanos, o no relativas, como suponen los antihumanos.

11. RESUMEN Y ADELANTO

El propósito de este capítulo ha sido proveer un análisis radicalmente antihumano de las razones normativas. De acuerdo con el análisis, decir que tenemos una razón normativa de hacer ϕ en ciertas circunstancias C es decir que, si fuésemos plenamente racionales, querríamos hacer ϕ en C . El análisis es radicalmente antihumano en cuatro aspectos.

Primero, hace que nuestras razones normativas sean el objeto de nuestras creencias, y también permite que nuestras creencias sobre nuestras razones normativas jueguen un rol propiamente causal en la producción de la acción. Segundo, nos brinda una perspectiva crítica aun sobre nuestros deseos no derivados, mostrándonos por qué podemos tener razones para deshacernos de ellos y razonar para adquirir otros deseos nuevos, no derivados, en cambio. Tercero, es una concepción no relativa de las razones normativas —lo que cuenta como una razón para ti en tus circunstancias también debe contar como una razón para mí en las mías si nuestras circunstancias son las mismas— y por lo tanto nuestras afirmaciones sobre nuestras razones normativas son imperativos categóricos y no imperativos hipotéticos. Y cuarto, nos fuerza a admitir que la epistemología de las afirmaciones de razón normativa es ella misma una cuestión social: todo lo demás igual, cada persona está tan bien situada como cualquier otra para llegar a una respuesta a la pregunta «¿qué razones normativas para actuar hay?». El análisis mismo no implica que haya razones normativas por supuesto —¿por qué habríamos de pensarlo así?—, pero sí nos permite empezar a preguntarnos si hay tales razones con una comprensión más clara de qué es lo que estamos buscando.

En el capítulo siguiente nuestro cómo estas consideraciones nos permiten resolver el problema moral del título del libro. También considero la cuestión sustantiva acerca de si hay razones normativas. Para anticipar: argumento que tenemos buena razón para creer que sí las hay.