

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

CAPÍTULO IV

LA TEORÍA HUMEANA DE LA MOTIVACIÓN

1. DOS PRINCIPIOS

Se ha argumentado que la teoría humeana de la motivación (de ahora en adelante «la teoría humeana») es un dogma en psicología filosófica, que el dogma es fundamentalmente incorrecto, y que por lo tanto debería ser reemplazado en psicología filosófica con una teoría de la motivación más plausible. Estoy pensando en particular en el trabajo reciente de Thomas NAGEL (1970), John McDOWELL (1978, 1979, 1981) y Mark PLATTS (1979, 1981). ¿Pero cuál es el dogma humeano?

Los humeanos aparecen comprometidos de hecho con dos afirmaciones sobre las razones motivadoras, una más débil y otra más fuerte. Sin embargo, no hay acuerdo entre los antihumeanos acerca de si la afirmación más débil o la más fuerte son ambas inaceptables o si es sólo la más fuerte la que deberíamos rechazar. De acuerdo con la afirmación más fuerte —la afirmación es, tal como yo la entiendo, crucial para la teoría humeana— la motivación tiene su fuente en la presencia de un deseo relevante y una creencia medio-fin. Esta afirmación puede presentarse de modo más formal con el siguiente principio (DAVIDSON, 1963):

- P1 R en t constituye una razón motivadora para el agente A para hacer ϕ si y sólo si hay algún Ψ tal que R en t consiste en un deseo apropiadamente relacionado de A hacia Ψ y una creencia de que si fuera a hacer ϕ entonces haría Ψ .

Decir que un deseo y una creencia deben estar «apropiadamente relacionados» es simplemente reconocer que para que un deseo y una creencia constituyan una razón motivacional el agente debe, por así decir, poner el deseo relevante y la creencia juntos (SMITH, 1988b).

Los antihumanos están unidos en su rechazo de P1. Sin embargo P1 implica un principio más débil según el cual la motivación requiere la *presencia* del deseo relevante y una creencia medio-fin, un principio que es a su vez consistente con la negación de la afirmación constitutiva hecha en P1. Así tenemos el siguiente principio más débil:

P2 El agente A en *t* tiene una razón motivacional para hacer ϕ sólo si hay algún Ψ tal que, en *t*, A desea hacer Ψ y cree que si fuera a hacer ϕ haría Ψ .

y los antihumanos no están todos unidos en su rechazo de P2. Así, por ejemplo, mientras que NAGEL (1970:29), y McDOWELL siguiéndolo (1978: 15), han argumentado que P2 es aceptable porque es consistente con la afirmación de que los deseos y las creencias medio-fin que deben estar presentes cuando hay motivación no son ellos mismos la *fuerza* de tal motivación — esto es, porque es consistente con la negación de la afirmación constitutiva hecha en P1 —, otros antihumanos, como PLATTS, han argumentado que P2 es inaceptable pues es o bien «fenomenológicamente falso [...] o completamente vacío» (1979: 256).

Estoy inclinado a acordar con los antihumanos que la teoría humeana es un dogma en la psicología filosófica, al menos en el sentido de que tanto P1 como P2 parecen encontrar un alto grado de aceptación no crítica. Sin embargo, a diferencia de los antihumanos, no creo que la teoría humeana, caracterizada en términos de P1, sea fundamentalmente incorrecta, y por lo tanto no pienso que P2 sea «completamente vacío».

Mi tarea en este capítulo es ofrecer un argumento explícito en favor de la teoría humeana y defenderla frente a objeciones ofrecidas por NAGEL, McDOWELL y PLATTS. Si el argumento ofrecido aquí es correcto entonces la teoría Humeana es la expresión de una verdad simple pero importante sobre la naturaleza de las razones motivacionales, una verdad que los antihumanos han fracasado en apreciar, o bien porque han fracasado en distinguir entre razones motivacionales de otro tipo de razones, o bien porque tienen una concepción inadecuada del deseo, o bien porque han pasado por alto las implicaciones del hecho de que las explicaciones de razón son teleológicas.

2. RAZONES MOTIVACIONALES VS. RAZONES NORMATIVAS

P1 es un principio que conecta razones *motivacionales* con deseos y creencias. Debemos empezar enfatizando este hecho, pues de lo contrario se consideraría simplemente implausible suponer que P1 provee condiciones individualmente necesarias o conjuntamente suficientes para que un estado constituya una razón motivacional.

Para ver esto, considérese los siguientes contraejemplos a la afirmación de que P1 provee condiciones necesarias:

- i) Supongamos que ahora deseo comprar un Picasso original, pero no creo que si fuera a comprar el cuadro que está delante de mí lo haría —supongamos que no creo que es un Picasso—. Si el cuadro delante de mí es de hecho un Picasso, entonces seguramente sería apropiado para que un extraño dijera que tengo una razón para comprar el cuadro delante de mí. Pero yo no tengo la creencia relevante.
- ii) Suponga que deliberadamente me paro sobre el pie de alguien para causarle daño; seguramente podemos imaginar que es apropiado que un extraño diga que tengo una razón para sacar el pie aun cuando yo no tenga el deseo relevante, y, de hecho, aun si deseo causar daño a esa persona.

Considérese ahora el siguiente contraejemplo a la afirmación de que P1 provee una condición suficiente:

- iii) Supongamos que ahora deseo tomar un *gin-tonic* y creo que puedo hacerlo mezclando el líquido que tengo delante de mí con agua tónica y tomándolo (WILLIAMS, 1980). Supongamos adicionalmente que esta creencia es falsa; que el líquido que tengo delante de mí es gasolina y no ginebra. Seguramente sería apropiado para un extraño decir que no tengo ninguna razón para mezclar este líquido con agua tónica y tomarlo. Sin embargo tengo tanto la creencia como el deseo relevantes.

¿Tenemos, en ejemplos como éstos, una objeción a P1 y por lo tanto a la teoría humeana? No. La razón de ello era quizás evidente desde el principio. La perspectiva del observador externo es irrelevante para los ejemplos.

Michael WOODS ha argumentado algún tiempo atrás que «el concepto de una razón para la acción está en el punto de intersección, por así decir, entre la teoría de la explicación de las acciones y la teoría de su justificación» (1972: 189). La idea de WOODS fue que las razones —*todas* las razones— tienen tanto una dimensión justificatoria como una explicativa. Dudo mu-

cho que esa afirmación sea verdadera, algo en favor de lo que argumentaré explícitamente después (capítulo V). Sin embargo Woods está seguramente en lo cierto al afirmar que nuestro concepto de una razón para la acción puede ser definido laxamente por estas dos dimensiones de explicación y justificación. De hecho, me parece a mí que trabajamos con dos conceptos bien diferentes de una razón para la acción dependiendo de si enfatizamos la dimensión explicativa y ponemos en segundo plano la justificatoria o viceversa. La afirmación «A tiene una razón para hacer φ » es por lo tanto ambigua. Puede ser una afirmación sobre una razón *motivacional* que A tiene, cuando enfatizamos la dimensión explicativa y ponemos en segundo plano la justificatoria, o una afirmación sobre una razón *normativa* que A tiene, cuando enfatizamos la dimensión justificatoria y ponemos en segundo plano la explicativa. Permítaseme decir algo sobre las similitudes y las diferencias entre estos dos tipos de razones.

Las razones motivacionales y normativas tienen algo en común en virtud de lo que ambas cuentan como razones. Pues citar una o la otra nos permite tornar la acción del agente en algo inteligible. Esto es esencial. Pues hay una conexión *a priori* entre citar las razones de un agente para actuar de cierto modo y hacer que ese actuar sea inteligible: esto es, especificando qué es lo que hay para decir en favor de actuar del modo en cuestión. En virtud de sus diferencias, sin embargo, las razones motivacionales y normativas hacen que las razones sean inteligibles por razones bien diferentes.

Decir que alguien tiene una razón normativa para hacer φ es decir que hay algún requerimiento normativo para que haga φ , y es por lo tanto decir que su hacer φ está justificado desde la perspectiva del sistema normativo que genera ese requerimiento. Como yo lo veo, y como argumentaré después, las razones normativas son mejor pensadas entonces como verdades: esto es, proposiciones de la forma general «que A haga φ es deseable o requerido» (capítulo V). Estas verdades pueden ser muchas y variadas, tan numerosas y variadas como los sistemas normativos que generen los requerimientos. Así, por ejemplo, puede haber razones normativas de la racionalidad, de la prudencia, de la moral, y quizás razones normativas de otros tipos. El número de diferentes tipos de razones normativas dependerá en última instancia de la medida en que podemos reducir un tipo de razón normativa a la otra. Por ejemplo, como argumentaré después, depende de si podemos reducir los requerimientos morales a requerimientos racionales.

Las razones motivacionales son, sin embargo, diferentes. El rasgo distintivo de una razón motivacional para hacer φ es que, en virtud de tener tal razón, un agente es un estado que es *explicativo* de su hacer φ , al menos todo lo demás igual —todo lo demás tiene que ser igual porque un agente puede tener una razón motivacional para hacer φ sin que la razón

sea concluyente—. Dado que un agente que tiene una razón motivacional para hacer φ es un estado que de este modo es potencialmente explicativo de su hacer φ es natural suponer que su razón motivacional es ella misma *psicológicamente real*. Pues parece ser parte de nuestro concepto de qué significa que las razones de un agente tengan el potencial de explicar su conducta que el tener esas razones es un hecho sobre ese agente. Que los objetivos que tales razones incorporan son *sus* objetivos. Por contraste con las razones normativas, entonces, que parecen ser *verdades* del tipo «es deseable o requerido que yo haga φ », las razones motivacionales parecen ser *estados psicológicos*, estados que juegan un cierto rol explicativo en la producción de la acción¹.

Si este modo de trazar la distinción entre razones motivacionales y normativas parece correcto, entonces estamos ahora en la posición de obtener las siguientes lecciones. Primero, las razones motivacionales y las razones normativas son diferentes *categorías*. Pues mientras que las razones motivacionales son estados psicológicos, las razones normativas son proposiciones de la forma general «que A haga φ es deseable o requerido». Y, segundo, hasta ahora no hay ninguna razón para suponer que hay una conexión general entre los dos tipos de razones. En función de todo lo que hemos visto hasta ahora, un agente puede por tanto tener una razón motivacional para hacer φ sin tener ninguna razón normativa para hacer φ , y puede tener una razón normativa para hacer φ sin tener ninguna razón motivacional para hacer φ . Dadas estas lecciones debemos enfatizar que P1 pretende brindar condiciones necesarias y suficientes para la existencia de razones motivacionales. P1 no dice nada acerca de si A tiene o no una razón normativa.

Ahora volvamos a los ejemplos que supuestamente traían problemas a P1 con esta distinción entre razones motivacionales y normativas en mente. Lo que vemos es que simplemente confirman la idea de que hay poca conexión entre los dos tipos de razones. Pues la perspectiva del extraño dirige nuestra atención a un requerimiento normativo relevante en cada caso. Los ejemplos por lo tanto no desafían en modo alguno la afirmación de que un agente tiene una razón motivacional si y sólo si tiene el deseo y la creencia relevantes.

¹ Las razones normativas son la temática de lo que Philip Pettit y yo, en otro lugar, llamamos la teoría de la racionalidad «de primer plano» (PETTIT y SMITH próximamente). Una teoría de la racionalidad de primer plano explicará en detalle las consideraciones que racionalmente justifican la acción. Pero nótese que podemos, bajo el nombre de «una teoría de la racionalidad», describir algo más también. Podemos describir la psicología del agente perfectamente racional; alguien que adopta tales razones de primer plano. PETTIT y yo la llamamos teoría de la racionalidad «de fondo». ¿Nos dirá una teoría de la racionalidad de fondo algo acerca de las razones motivacionales que tiene un agente? Solamente si el agente en cuestión es perfectamente racional. Para mayor abundamiento del tema, véase el capítulo V.

En i), por ejemplo, es suficiente para la verdad de la afirmación de que tengo una razón para comprar el cuadro delante de mí que hay un requerimiento — en este caso, en sentido amplio, un requerimiento de la razón — para que compre el cuadro delante de mí². Pues yo pienso que comprar un Picasso es deseable y el cuadro delante de mí es de hecho un Picasso. Pero la existencia de tal razón normativa no es suficiente para que yo tenga una razón motivacional para comprar el cuadro delante de mí. Esto es así porque, ya que no creo que el cuadro sea un Picasso, no estoy en un estado que es potencialmente explicativo de mi acción de comprarlo. (Estoy por supuesto en un estado psicológico que sea potencialmente explicativo de que compre un Picasso, pues deseo comprar un Picasso y creo que, si fuera a comprar un Picasso, compraría un Picasso. Pero esto no es tener una razón motivacional para comprar el cuadro delante de mí. Es más bien una razón motivacional para comprar un Picasso. El humeano por lo tanto considerará el deseo de un agente de hacer ϕ , junto con la creencia trivial de que si fuera a hacer ϕ haría ϕ , como el caso límite de tener una razón motivacional para hacer ϕ). El ejemplo entonces de ningún modo socava la necesidad de tener una creencia medio-fin para tener una razón motivacional. Simplemente dirigen nuestra atención al hecho de que puedo tener o no una razón motivacional para hacer lo que racionalmente me es requerido.

Un punto similar se aplica a ii), el caso en que tengo una razón para dejar de pisar a alguien cuando le causo dolor. Pues es suficiente para la verdad de la afirmación de que tengo una razón para no pisarlo que existe un requerimiento — en este caso moral — de que no le cause dolor, y de que, en las presentes circunstancias, para cumplir con ese requerimiento tengo que sacarle el pie de encima. Pero, de nuevo, la mera existencia de esta razón normativa es consistente con la afirmación de que no estoy en un estado que es potencialmente explicativo de mi conducta. De hecho, adviértase que esto debe ser concedido aun si el argumento de este libro como un todo es correcto y las razones morales son ellas mismas requerimientos racionales sobre la acción. Pues, como ya hemos visto, los requerimientos racionales son a su vez simplemente razones normativas, que por lo tanto pueden existir en la ausencia de razones motivacionales. Así, este tipo de ejemplos no muestra por sí mismo que tener un deseo no sea una condición necesaria para tener una razón motivacional.

² Al sugerir que el requerimiento es un requerimiento de racionalidad «en sentido amplio» estoy siguiendo a WILLIAMS (1980: 102-103) y PARFIT (1984: nota 2a hasta parte I; 117-120). Ambos piensan que una teoría de la racionalidad práctica nos diría qué razón tiene un agente para actuar, y ambos piensan que esto, a su turno, depende, inter alia, de cuál es el caso, antes de que, simplemente, lo que el agente cree que es el caso.

Consideremos ahora iii), el contraejemplo para la condición suficiente. ¿En qué sentido no tengo una razón para mezclar el líquido delante de mí con agua tónica y tomarlo? Otra vez parece que la razón es normativa más que motivacional. Pues ciertamente estoy en un estado que es potencialmente explicativo de que mezcle el líquido delante de mí con agua tónica y lo tome. Después de todo, deseo tomar un gin-tonic y creo que el líquido delante de mí es ginebra. Más aún, si explicásemos que yo haya hecho esto citando ese deseo y esa creencia mi acción sería ciertamente inteligible. Pues podría ser visto como apuntando a ese objetivo que tengo. Por lo tanto, aunque hay un sentido en que no tengo una razón para mezclar el líquido delante de mí con agua tónica y tomarlo — la prudencia, por ejemplo, me aconsejaría en contra — la razón que no tengo es una razón normativa. Parece enteramente correcto suponer que ahora tengo una razón motivacional para hacer justamente eso. Por lo tanto no tenemos aquí un contraejemplo a la afirmación de que P1 provee de una condición suficiente para que un estado constituya una razón motivacional.

Para resumir: a luz de la distinción de estas razones motivacionales y normativas, hemos visto que la teoría humeana es una teoría sobre la naturaleza de las razones motivacionales. La teoría humeana todavía puede ser falsa. Pero no se muestra que sea falsa simplemente mostrando que P1 fracasa al proponer condiciones necesarias y suficientes para la existencia de una variedad de razones normativas.

3. UNA OBJECION PRELIMINAR DE NAGEL

Adviértase que las dos afirmaciones en favor de las cuales acabo de argumentar — que hay una distinción entre razones motivacionales y normativas, y que la teoría humeana es una teoría sobre las primeras y no sobre las segundas — son suficientes para socavar uno de los principales argumentos de THOMAS NAGEL contra la teoría humeana en *The Possibility of Altruism* (1970). La objeción de NAGEL se centra en la explicación humeana de la motivación prudencial.

La motivación prudencial es posible sólo si el reconocimiento por parte de un agente del hecho de que tendrá un deseo de hacer ϕ en el futuro le da, de algún modo, una razón ahora para adoptar los pasos para promover su hacer ϕ en el futuro. La tarea de explicar esta posibilidad adopta una forma particular para el humeano. Pues, como hemos visto, él afirma que tener una razón motivacional para hacer ϕ ahora requiere desear ϕ ahora. Debe por lo tanto explicar cómo el reconocimiento de un agente de que deseará ϕ en el futuro le da un deseo presente de promover ϕ entonces. La respuesta del

humano es bastante predicable. Dice que los agentes que están motivados por consideraciones prudenciales tienen cada uno un deseo presente general de promover sus intereses futuros.

NAGEL ofrece la siguiente objeción a la explicación de la prudencia humana.

Los dos rasgos del sistema que objeto son *a*) que no permite que la expectativa de una razón futura nos provea ella misma ninguna razón para la acción presente, y *b*) que permite que un deseo presente por un objeto futuro provea por sí mismo una razón para la acción presente en persecución de ese objeto (1970: 39).

Así, tal como él señala, las siguientes son posibilidades en la teoría humana.

Primero, dado que cualquier deseo de un objeto futuro provee una base para razones en favor de hacer aquello que promoverá ese objeto, puede suceder que ahora desee algo para el futuro que no espero ni esperaré desear entonces, algo que no creo que haya ninguna razón en el futuro para promover. En consecuencia puedo tener una razón ahora para prepararme a hacer aquello que sé que no tendré ninguna razón para hacer cuando llegue el momento.

Segundo, supóngase que espero ser asaltado por un deseo en el futuro: entonces debo reconocer que en el futuro tendré una razón *prima facie* para hacer lo que el deseo indica. Pero esta razón no aparece ahora, y no puede por sí misma aplicarse derivativamente a ningún medio disponible presente para satisfacer el deseo futuro. Así, en la ausencia de algún deseo relevante adicional en el presente, puedo no tener una razón para prepararme a hacer aquello que yo sé que tendré razón para hacer mañana (1970: 39-40).

La respuesta que NAGEL pretende que nosotros demos, al confrontar estos ejemplos, es que en el primer caso no tengo una razón para promover el objeto futuro a pesar de mi deseo presente, y que en el segundo tengo una razón para promover el objeto de mi deseo futuro a pesar de no tener el deseo relevante presente.

Esa es la respuesta y me parece que es la respuesta correcta. Pero no creo que este hecho cuente en contra de la teoría humana. Para mostrar por qué, considérese el resumen de la objeción que el propio NAGEL ofrece.

Un sistema con consecuencias como las mencionadas no sólo fracasa en requerir la consistencia más elemental en la conducta a lo largo del tiempo, sino que de hecho aumenta las posibilidades de conflicto al basar los planes de un individuo contra sí mismo en el futuro en el aparato de la racionalidad.

Hay dificultades formales y extremadamente generales con este sistema, pues conciernen a la relación entre qué es racional y qué será racional sin importar qué fuente de las razones es operativa (1970: 40-41).

Así, si aceptamos el propio diagnóstico de NAGEL sobre la respuesta a estos ejemplos —y pienso que deberíamos— es evidente que los ejemplos como estos fracasan siquiera en tocar al humano. Pues, para tomar simplemente el primero (el problema con el segundo es prácticamente el mismo), la objeción de NAGEL contra la afirmación humana de que un agente puede tener una razón motivacional ahora para promover su acción de hacer ϕ en el futuro, a pesar de que cree que no tendrá ninguna razón motivacional para hacer ϕ entonces, es que sería ahora irracional hacerlo. Que ahora no tiene ninguna razón desde la perspectiva de la racionalidad para hacerlo (véase también NAGEL, 1970: 64-65). Pero esto es confundir la afirmación de que un agente tiene una razón motivacional para hacer ϕ con la afirmación de que tiene una razón normativa desde la perspectiva de la racionalidad para hacer ϕ . El Humano solamente está haciendo la primera afirmación, no la segunda.

Más aún, si NAGEL está en lo cierto al decir que es irracional para un agente promover su hacer ϕ en el futuro si cree que entonces no tendrá ninguna razón motivacional para hacer ϕ , entonces el teórico humano de la motivación puede aceptar esto en sus propios términos, quizás afirmando que, ya que no hay ninguna razón normativa de la racionalidad para que los agentes promuevan sus intereses futuros, así también la teoría de la racionalidad requiere que los agentes tengan el deseo de promover sus intereses futuros (PARFIT, 1984: 131-136). Pues, de modo importante, la tarea de construir una teoría de las razones motivacionales y una teoría de las razones normativas son diferentes tareas³.

Por lo tanto no veo que la racionalidad de la prudencia cree una especial dificultad a la teoría humana de la motivación. Pensarlo así requiere confundir la teoría humana de la motivación con una teoría de la racionalidad.

³ Nótese que nosotros podemos distinguir las dos afirmaciones que son hechas aquí en términos de la distinción efectuada en la nota 1. Decir que los agentes son requeridos racionalmente para promover sus intereses futuros, o que hay una razón normativa de racionalidad para hacerlo, es hacer una afirmación acerca de la teoría de la racionalidad de primer plano (una afirmación sobre las consideraciones que racionalmente justifican la acción), mientras que decir que a los agentes se les requiere racionalmente tener un deseo de promover sus futuros intereses es hacer una afirmación sobre la teoría de la racionalidad de fondo (una afirmación sobre la psicología del perfecto agente racional). Relacionado con ello, nótese que un humano que piensa que los agentes tienen y actúan según un deseo para promover sus intereses futuros no necesitan, y quizás no deberían, pensar que tienen que apelar a ese deseo para justificar sus acciones (PETTIT y SMITH, 1990: 589-591).

4. ¿POR QUÉ DEBERÍAMOS CREER EN LA TEORÍA HUMEANA?

Hemos visto que no encontraremos una refutación fácil de P1 —la afirmación de que las razones motivacionales están constituidas por deseos apropiadamente relacionados con creencias medio-fin— a través de la reflexión sobre aquellos casos en que ordinariamente diríamos de alguien que tiene una razón para hacer ϕ . ¿Pero podemos encontrar una razón general para de hecho creer en esa afirmación?

John McDOWELL ha intentado diagnosticar el compromiso con la teoría Humeana (1981). Él empieza aislando lo que considera el rasgo distintivo de la teoría, esto es, que «citar una actitud proposicional cognitiva», esto es, una creencia, «es brindar como máximo una especificación parcial de una razón para actuar. Para ser completamente explícito, uno debería agregar la mención a algo no cognitivo, un estado de la voluntad o un evento volicional» (1981: 155): esto es, en los términos en que lo hemos puesto, un deseo. McDOWELL continúa:

Sospecho que una razón por la cual gente encuentra [...] [que esta afirmación] [...] es obvia descansa en su adherencia no explícita a una concepción cuasihidráulica de cómo las explicaciones de razón dan cuenta de la acción. La voluntad es presentada como la fuente de fuerzas que desembocan en la conducta que tales explicaciones explican. Esta idea me parece a mí una concepción radicalmente equivocada del tipo de explicación que la explicación de razón es, pero no es mi preocupación presente (1981: 155).

¿Pero qué es exactamente una concepción «cuasihidráulica» de la explicación de razón?

De acuerdo con la concepción cuasihidráulica de las explicaciones de razón, la mente es un territorio donde varias fuerzas o presiones (McDOWELL dice «deseos») se canalizan (a través de creencias) en ciertas direcciones y en última instancia se combinan para producir una fuerza o presión resultante (una acción). Quizás McDOWELL está en lo cierto al afirmar que el compromiso de esa concepción de la explicación de razón ha explicado, en el pasado, por qué la gente acepta la teoría humeana. Pero, aun si acordamos con McDOWELL respecto de que esa concepción está «radicalmente equivocada», no es claro cómo esto podría arrojar alguna duda en la teoría humeana. De hecho, suponer que lo hace suena mucho como la afirmación del consecuente. Lo que arrojaría duda en la concepción Humeana no es el hecho de que alguna gente que la acepta lo hace o lo ha hecho porque se adhiere a una concepción equivocada de la explicación de razón, sino más bien que adherirse a la teoría humeana misma *requiere* la adopción de esa

concepción equivocada de la explicación de razón. Pero, como argumentaré ahora, esta última afirmación es manifiestamente implausible. Pues lejos de estar comprometido a una concepción cuasihidráulica de las explicaciones de razón, el humeano no está ni siquiera comprometido a una concepción causal de las explicaciones de razón. Dado que una concepción cuasihidráulica es una forma de concepción causal, se sigue que la teoría humeana simplemente no está comprometida con una concepción cuasihidráulica de las explicaciones de razón.

Para ver esto, considérese el siguiente argumento que los teóricos causales de hecho dan en favor de la concepción causal de las explicaciones de razón (DAVIDSON, 1963: 8-11). «¿Cuál es el rasgo que hace la diferencia entre el caso de un agente que hace ϕ y tiene una razón para hacer ϕ y el caso en que un agente hace ϕ *porque* tiene una razón para hacer ϕ ? La única respuesta iluminadora disponible» afirma el teórico causal, «es que las razones en el segundo caso —el caso en que la razón explica la acción— *causa* que haga ϕ ». Ahora bien, no estoy aquí preocupado por los méritos de este argumento. Más bien quiero enfatizar el hecho de que el argumento no formula ninguna suposición sustancial sobre la naturaleza de las razones motivacionales. Así, lo único que el argumento nos dice es que las razones motivacionales pueden estar constituidas por creencias o por deseos o por algún otro tipo de estado psicológico. Sin embargo, si esto es correcto, se sigue que deberíamos ser capaces de aceptar o rechazar este argumento independientemente de nuestras visiones relativas a la teoría de la motivación.

Supongamos, por ejemplo, que rechazamos el argumento de DAVIDSON. Puede parecer evidente por qué lo haríamos. Pues podríamos pensar que hay algo más básico y aun así iluminador para decir sobre el «porque» en la frase «ella hace ϕ porque tiene una razón para hacer ϕ ». Pues, podríamos decir, el «porque» aquí señala la disponibilidad de una explicación *teleológica* de la acción del agente en términos de su razón para hacer ϕ , una explicación que no está necesariamente disponible cuando todo lo que sabemos es que el agente tiene una razón para hacer ϕ y tiene una razón para hacer ϕ (MCGINN, 1982, 99-100; WILSON, 1985: 31-32). Más aún, podríamos decir, es controvertido, y ciertamente no está forzado simplemente por el hecho de que tenemos que admitir la diferencia entre «ella hace ϕ y tiene una razón para hacer ϕ » y «ella hace ϕ porque tiene una razón para hacer ϕ », suponer que esta explicación teleológica puede después ser presentada en términos *causales*. De modo importante, sin embargo, adviértase que esto es algo que tanto el teórico humeano como el antihumeano pueden *igualmente* decir en respuesta al argumento que se acaba de dar en favor de la teoría causal. Pues no hay nada en el debate sobre la teoría de la motivación *per se* que nos fuerce a suponer que las razones motivacionales son causalmente, más que teleológicamente, explicativas. De hecho, la teoría causal es mejor pensada

como una interpretación, quizás opcional, de la afirmación de que las explicaciones de razón son teleológicas.

Ahora supongamos que aceptamos el argumento de DAVIDSON; que pensamos que la mejor interpretación del carácter teleológico de las explicaciones de razón es una interpretación en términos causales. Entonces, con seguridad debemos concebir algunos estados psicológicos que posean el poder causal para producir la conducta. Pero, de nuevo, no necesitamos pensar que los deseos son el único estado psicológico que posee ese poder causal. Podríamos pensar en cambio que sólo ciertas creencias, o algún otro tipo completamente distinto de estado psicológico, poseen ese poder causal. Si pensamos que los deseos o las creencias o algún otro estado psicológico tienen poder causal, por supuesto, dependerá de si aceptamos una teoría de la motivación humeana o antihumeana. Pero ahora esto debe ser visto como algo completamente independiente de si, además, deberíamos aceptar o no una interpretación causal del carácter teleológico de las explicaciones de razón.

El resultado parece entonces ser que los teóricos de la motivación humeana y antihumeana están involucrados en un debate que es no sólo independiente sino más fundamental que el debate sobre la cuestión de si las explicaciones de razón son una especie de explicación cuasihidráulicas, o más simplemente causal. En definitiva la diferencia es ésta. El teórico causal y el no causal pueden ambos aceptar que las explicaciones de razón son explicaciones teleológicas sin investigar adicionalmente la cuestión de qué hay en la naturaleza de las razones que hace posible que las explicaciones de razón sean explicaciones teleológicas — es decir, explicaciones que explican haciendo inteligible aquello que explican en términos de la persecución de un objetivo —. Pues su desacuerdo es relativo a la cuestión posterior acerca de si las explicaciones teleológicas son ellas mismas una especie de explicación causal, un desacuerdo que, como he sugerido, atraviesa los desacuerdos relativos a la naturaleza de las razones mismas. Pero, como yo lo veo, el humeano y el antihumeano están precisamente involucrados en una disputa relativa acerca de cuál es la naturaleza de las razones que hace posible que las explicaciones de razón sean explicaciones teleológicas.

Si esto es correcto, entonces parece que sólo habrá una razón para creer la teoría humeana, si hay alguna razón para creerla, y esta razón es que sólo esta teoría es capaz de dar sentido a la motivación como la persecución de un objetivo. En lo que sigue argumentaré que éste es de hecho el caso. Procederé concentrándome en dos diferentes concepciones del deseo. Una de ellas no brinda ningún apoyo a la teoría humeana. Argumento que esta concepción es de todos modos implausible. Quizás de modo no sorprendente es la concepción del deseo adoptada por muchos oponentes de la teoría humeana.

Hay, sin embargo, una concepción alternativa y más plausible. Esta concepción nos permite ver que los deseos deben ser constitutivos de las razones motivacionales dado que las razones motivacionales de un agente deben estar ellas mismas constituidas por que ellos tengan ciertos objetivos.

5. DESEOS Y FENOMENOLOGÍA

¿Qué es un deseo?

De acuerdo con HUME, los deseos son especies de las pasiones, y las pasiones son, a su vez, un cierto tipo de sentimiento. HUME parece decir que esto es así no sólo en el sentido trivial de que las pasiones son una especie de la percepción y las percepciones son un tipo de sentimiento (1888: 190). Más bien él sugiere que cuando deseamos algo «sentimos una [...] emoción de aversión o propensidad» (1888: 414). Su visión entonces parece ser que estamos «directamente conscientes» de la presencia de los deseos que tenemos; que los deseos son ellos mismos fenomenológicamente salientes (STROUD, 1977: 163).

La sugerencia de HUME no está enteramente descarriada. Hay, después de todo, algo así como una fenomenología del deseo, como por ejemplo, para usar uno de los ejemplos del propio HUME, «cuando estoy con rabia estoy poseído por la pasión» (1888: 416). Esto es, podemos acordar con HUME que, en ocasiones, cuando tengo un deseo, tengo ciertos sentimientos psicológicos, análogos a las sensaciones corporales. Esto puede sugerir un modo de elaborar la visión humeana. Pues si consideramos su sugerencia de que todos los deseos son conocidos por el modo en que nos hacen sentir, entonces, en un intento por explicar por qué ello es así, podríamos ser llevados a identificar los deseos con tales sentimientos psicológicos. Y esto a su vez nos podría llevar a adoptar lo que llamaré «la concepción fenomenológica fuerte» de los deseos: la visión de que los deseos son, como las sensaciones, simple y esencialmente estados que tienen cierto contenido fenomenológico.

Quizás de modo no sorprendente, pienso que la concepción fenomenológica fuerte del deseo debe ser rechazada. Pues me parece que no hay modo en que esa concepción pueda ser hecha compatible con una epistemología plausible del deseo.

Ahora bien, podría pensarse que esta objeción ni siquiera despega del piso. La concepción fenomenológica fuerte del deseo hace que la epistemología del deseo sea no problemática, podría decirse, pues en tal concepción la epistemología del deseo se convierte simplemente en la epistemología de la sensación. Así como es plausible afirmar que un sujeto siente dolor si y

sólo si cree que tiene dolor -pues suponemos que un sujeto está en un estado con cierto contenido fenomenológico si y sólo si cree que está en ese estado con ese contenido— así, si pensamos los deseos en función del modelo de las sensaciones, es posible afirmar que un sujeto desea que ϕ si y sólo si cree que desea que ϕ .

¿Qué anda exactamente mal con esta línea de pensamiento? Hay dos cosas que andan mal. Para empezar, y de modo más directo, el principio «los sujetos desean que ϕ si y sólo si creen que desean que ϕ » es simplemente falso. Argumento mediante contra-ejemplos.

Supongamos que cada día en su camino al trabajo John compra un periódico en cierto quiosco. Sin embargo, tiene que desviarse de su camino para hacerlo, y por ninguna buena razón aparente. El periódico que compra está disponible en otros quioscos en su ruta directa al trabajo, no hay diferencia en el precio, o en la condición de los periódicos entre los diferentes quioscos, etc. Hay, sin embargo, la siguiente diferencia. Detrás del quiosco en el que John compra su periódico, hay espejos dispuestos de tal modo que cualquiera que compra un periódico no puede evitar mirarse a sí mismo. Supongamos, sin embargo, que si fuera sugerido a John que la razón por la que compra su periódico en ese quiosco es que quiere mirar su propio reflejo, él lo negaría vehementemente. Y no parecería a John que se estuviera autoengañando al negar eso. Sin embargo, finalmente, supongamos que si los espejos fueran removidos del quiosco, su preferencia por ese quiosco desaparecería.

Si todo esto fuera el caso, ¿no sería plausible suponer que John en realidad desea comprar el periódico en el quiosco en el que puede mirar su propio reflejo; que, quizás, tiene una tendencia narcisista y que comprar el periódico en ese quiosco le permite complacerse en su camino al trabajo? ¿Y no sería también plausible suponer que no cree que esto es así, dada, desde su punto de vista, su negativa sincera? Si esto es aceptado, entonces tenemos una razón para rechazar el principio de izquierda a derecha: esto es, tenemos razón para negar que si un sujeto desea que ϕ entonces cree que desea que ϕ ⁴.

Considérese otro ejemplo. Supongamos que John profesa que uno de sus deseos más fundamentales es ser un gran músico. Sin embargo, su ma-

⁴ De hecho, me parece que hay contraejemplos más mundanos al principio considerado de izquierda a derecha. Considérese casos en que usted va al refrigerador convencido de que hay algo en particular que usted quiere, aunque no está muy seguro de qué es. Después, al mirar el contenido del refrigerador, repetinamente (podría decirse) «se da cuenta qué era lo que quería todo este tiempo». Si queremos respetar esta descripción de sentido común de esas situaciones entonces deberíamos rechazar el principio de izquierda a derecha.

dre siempre le ha insistido sobre el valor de la música. Ella es una fanática con grandes esperanzas para la carrera de su hijo como músico, esperanzas tan grandes que estaría completamente desilusionada si fuera menos que un músico excelente, y mucho más si fuera a abandonar la música. Más aún, John admite que tiene un gran deseo de no desilusionarla, aunque, si se le preguntara, negaría que esto de algún modo explique sus esfuerzos en la persecución de la excelencia en la música. Sin embargo, ahora supongamos que la madre de John muere, y que después de su muerte John encuentra que todo su interés en la música desaparece. Abandona su carrera como músico e inicia una carrera diferente. En esas circunstancias, ¿no sería plausible suponer que John simplemente estaba equivocado sobre lo que originalmente quería hacer y que, a pesar del hecho de que creía que lograr la excelencia en la música era un deseo fundamental suyo, en realidad nunca lo fue? Si es así, entonces tenemos una razón para creer que el principio es falso de derecha a izquierda también: esto es, tenemos razón para negar que si un sujeto cree que tiene una razón para hacer ϕ entonces desea que ϕ .

Si esto es aceptado, entonces cualquier concepción de los deseos que implique que un sujeto desea que ϕ si y sólo si cree que desea que ϕ es una concepción que debe rechazarse. Tiene que rechazarse por la simple razón de que la epistemología que provee es implausible. Así, si la concepción fenomenológica implica ese principio, debe rechazarse. Y esto a su vez nos enseña la siguiente valiosa lección. Es un requisito de adecuación de cualquier concepción del deseo que la epistemología del deseo que recomienda permita que los sujetos sean *fallibles* sobre los deseos que tienen.

Dije al principio que hay dos líneas de objeción a la concepción fenomenológicamente fuerte del deseo. La segunda es aún más poderosa que la primera. Supongamos que concedemos que los deseos son como las sensaciones en el sentido de que tienen esencialmente un contenido fenomenológico. Aun así, debemos aceptar que difieren de las sensaciones en que, además, tienen contenido proposicional (PLATTS, 1981: 74-77). La adscripción de deseos, a diferencia de la adscripción de sensaciones, puede ser hecha de la forma «A desea que p », donde « p » es una oración. Por lo tanto, mientras que el deseo de A de hacer ϕ puede ser adscrito a A en la forma «A desea que ϕ », el dolor de A no puede ser adscrito a A en la forma «A duele que p ».

Es por lo tanto ambiguo afirmar que la epistemología del deseo es como la epistemología de la sensación. Con certeza, si los deseos son estados esencialmente fenomenológicos entonces la epistemología del *contenido fenomenológico* del deseo puede estar basada en la epistemología de la sensación. ¿Pero qué decir sobre la epistemología del *contenido proposicional* del deseo? Ésta no puede estar basada en la epistemología de la sensación en absoluto, pues las sensaciones no tienen contenido proposicional.

Por lo tanto se sigue que tenemos una razón incluso más fuerte para rechazar la concepción fenomenológica fuerte del deseo. Pues, de acuerdo con esta concepción, no hay diferencia entre deseos y sensaciones. Cada uno es un estado que simple y esencialmente tiene contenido fenomenológico. La concepción fenomenológicamente fuerte de los deseos es por lo tanto incapaz de dar cuenta del hecho de que los deseos tienen contenido proposicional. Es poco sorprendente que no pueda proveer una epistemología plausible del contenido proposicional de los deseos.

Sospecho que, por esta misma razón, alguien puede pensar que la concepción fenomenológica fuerte del deseo era un hombre de paja desde el principio. Pero adviértase que, una vez que están firmes nuestras objeciones a la concepción fenomenológica fuerte, estamos ahora en una posición para argumentar en contra de *todas* las versiones de la concepción fenomenológica, aun las concepciones plausibles más débiles de acuerdo con las cuales los deseos son como las sensaciones en el sentido de que tienen un contenido fenomenológico esencial, pero difieren de las sensaciones en el sentido de que tienen contenido proposicional también. Pues ahora podemos decir lo siguiente sobre todas esas concepciones. De ningún modo contribuyen a nuestra comprensión de qué es un deseo en tanto estado con contenido proposicional, pues no pueden explicar cómo es que los deseos tienen contenido proposicional. Por lo tanto no explican de modo alguno la epistemología del contenido proposicional del deseo. Y por lo tanto requieren complementarse con alguna teoría independiente y con sustento propio de qué es un deseo que explique cómo es que los deseos tienen contenido proposicional y cómo es que tenemos un conocimiento falible de qué es lo que deseamos.

La cuestión que inmediatamente se presenta respecto de las concepciones fenomenológicas más débiles es entonces por qué deberíamos creer que una concepción de ese tipo es verdadera. La única respuesta disponible es que la concepción fenomenológica da cuenta de la fenomenología del deseo. ¿Pero es esta respuesta plausible? ¿Realmente creemos que los deseos son estados de cosas que tienen cualidades fenomenológicas esencialmente? Esto es, ¿creemos acaso que si no hay algo así como un deseo, en cierto tiempo, entonces no se tiene ese estado en ese tiempo?

Debería decir, al menos, hasta dónde llega la opinión del sentido común —¿y qué otra cosa deberíamos considerar al formular una concepción filosófica de los estados psicológicos de la gente común?— que evidentemente no tenemos esa creencia. Pensemos, por ejemplo, en lo que ordinariamente consideraríamos un deseo de largo término: digamos, el deseo del padre de que a su hijo le vaya bien. Un padre quizás sienta el pinchazo de este deseo de tiempo en tiempo, en momentos de reflexión sobre su vulnerabilidad,

digamos. Pero esas ocasiones no son la norma. Y sin embargo, con certeza no pensaríamos que ha perdido su deseo en esos períodos donde no tiene su sentimiento. O consideremos casos más mundanos en los que, como decimos ordinariamente, deseo cruzar la calle y lo hago, o en que deseo escribir algo y lo escribo. Como STROUD señala, en esos casos «es difícil creer que estoy sobrepasado por la emoción [...] ciertamente no soy consciente de ninguna emoción o pasión que me empuje a actuar»; más bien «parecen el modelo mismo de acción desapasionada y fría» (STROUD, 1977: 163). Sin embargo estaría groseramente en contra de nuestra opinión de sentido común concluir que simplemente porque no advierto introspectivamente la presencia de deseos en esos casos entonces me atribuyo incorrectamente deseos a mí mismo —que cruzo la calle y escribo cosas ¡aunque no lo quiero!—.

Por supuesto, si pensáramos que no queda nada del deseo si no es sentido entonces podríamos, en nuestro rol de teóricos filosóficos, sentirnos obligados a concluir que algunas de nuestras atribuciones de deseos ordinarias están equivocadas, y por lo tanto sentirnos obligados a revisar nuestras opiniones ordinarias en favor de la concepción fenomenológica. Pero dado que hemos visto que esa concepción es incapaz de proveer una teoría de qué es un deseo como estado con contenido proposicional, no deberíamos sentir ninguna presión en nuestro rol de teórico filosófico. Más bien deberíamos conceder que podemos tener un deseo y no tener el sentimiento correspondiente.

Esto es significativo. Pues muchos antihumeanos parecen trabajar con una concepción fenomenológica del deseo, y después usan el hecho de que no advertimos interiormente la presencia de deseos cuando hay motivación en contra de la teoría humeana. Considérese, por ejemplo, el siguiente argumento de Mark PLATTS:

La premisa crucial [...] es la afirmación de que cualquier especificación completa de una razón para la acción, si es una razón para la acción para el agente potencial, debe hacer referencia a los deseos del agente. A primera vista, parece un hecho doloroso de la vida moral que esta premisa sea falsa. Llevamos adelante muchas acciones intencionales en esa vida que aparentemente no deseamos llevar adelante. Una descripción mejor de esos casos parece ser que llevamos a cabo esas acciones porque las consideramos deseables. La dificultad de gran parte de la vida moral emerge entonces como una consecuencia del hecho aparente de que desear algo y considerarlo deseable son cosas distintas e independientes. La premisa puede, por supuesto, considerarse verdadera simplemente afirmando que, cuando actuamos porque pensamos que algo es deseable, de hecho lo deseamos. Pero esto es o bien fenomenológicamente falso, pues no hay nada en nuestra vida interior que se corresponda con el deseo supuesto, o completamente vacío, de modo que ni el contenido ni la motivación se deban a la existencia del supuesto deseo.

Nada más que confusión (y aburrimiento) se desprende de tratar a los deseos como un captúralo-todo mental (1979: 256).

Así, de acuerdo con PLATTS, el humeano puede sostener que cuando actuamos creyendo que algo es deseable lo deseamos, pero si lo hace, está frente a los cuernos de un dilema. Pero consideremos los cuernos del dilema de PLATTS.

Si no hay ninguna razón por la cual un teórico debería aceptar una concepción fenomenológica del deseo, tal como hemos visto que es el caso, entonces difícilmente pueda hacer una objeción a la teoría humeana que seamos incapaces de advertir introspectivamente la presencia de todos y cada uno de los deseos que dice que tenemos. Entonces así no deberíamos adjudicar forzosamente al humeano la falsedad fenomenológica en el primer cuerno. PLATTS podría estar de acuerdo. Pero, diría yo, esto simplemente fuerza al humeano a admitir el segundo cuerno del dilema.

Aquí PLATTS afirma que si no aceptamos la concepción fenomenológica del deseo, entonces suponer la existencia de un deseo debe ser «completamente vacío» o sin «contenido». Pero, como hemos visto, dada la suposición de que los deseos son estados con contenido proposicional, una suposición que debe ser aceptada aun por aquellos que adoptan la concepción fenomenológica, el único modo de que podemos dar contenido al concepto de ese estado, y por lo tanto a la existencia del supuesto deseo, es precisamente *vía* alguna concepción no fenomenológica independiente y aceptable en sus propios términos. Por tanto, las concepciones no fenomenológicas, lejos de hacer que la adscripción de deseo sea «completamente vacía», son las únicas que hacen que la adscripción del deseo con contenido proposicional sea posible.

De hecho, aun John McDOWELL, quien rechaza la concepción fenomenológica, implícitamente adjudica esa concepción al humeano cuando argumenta en contra de él⁵. Esto emerge en la defensa de McDowell de su propia visión de que el agente virtuoso puede estar motivado por su concepción de la situación en la que se encuentra, algo que, de acuerdo a McDowell, podría ser considerado apropiadamente un estado cognitivo. Él supone correctamente que el humeano replicaría que, si alguien con tal concepción está de hecho motivado, entonces lograr que tenga esa concepción debe involucrar lograr tener cierto deseo. Pero luego él interpreta esa sugerencia como

⁵ Yo digo que McDOWELL rechaza una concepción fenomenológica. Él en ningún lado dice que rechaza tal concepción. Pero debe ser así, dado que él piensa que «los deseos atribuidos consecucionalmente son efectivamente deseos» (1978: 25). Lo importante de los deseos atribuidos consecucionalmente es que posiblemente no hay una base fenomenológica para su adscripción. Discuto esta idea hacia el final del presente capítulo.

«Véalo de este modo», es realmente una invitación implícita a sentir, más allá de nuestra propia visión de los hechos, un deseo que se combinará con la creencia de uno para recomendar actuar del modo apropiado (1978: 22).

Y correctamente rechaza esta sugerencia. Pero, si el humeano no tiene que aceptar la concepción fenomenológica del deseo, ¿por qué darle esa interpretación a su réplica? Con seguridad, lograr que alguien tenga cierta visión de los hechos que lo llevará a actuar de cierto modo puede no involucrar lograr que *sienta* un deseo, pero puede involucrar lograr que *tenga* de todos modos un deseo.

6. DESEOS, DIRECCIONES DE AJUSTE Y DISPOSICIONES

Hemos visto que debe haber una alternativa a las concepciones fenomenológicas del deseo, una alternativa que nos permita dar sentido a los deseos como estados con contenidos proposicionales y que por tanto nos permita dar sentido a nuestras atribuciones ordinarias de deseo. ¿Pero cuál es esta alternativa?

De modo sorprendente, PLATTS mismo describe la alternativa que yo favorezco, una sugerencia sobre la diferencia entre creencia y deseo que atribuye a ANSCOMBE (1957). El resumen de PLATTS es tan sucinto y hace que la idea suene tan plausible que simplemente lo voy a citar. Debajo considero por qué PLATTS mismo está tan poco inclinado a aceptar la idea.

La señora Anscombe, en su trabajo sobre la intención, ha trazado una distinción gruesa entre dos tipos de estados mentales, la creencia empírica es el ejemplo principal de un tipo y el deseo el ejemplo principal del otro [...] La distinción se propone en términos de direcciones de ajuste de los estados mentales y el mundo. Las creencias apuntan a ser verdaderas, y su ser verdaderas depende de que se ajusten al mundo. La falsedad es un fracaso decisivo en una creencia, y las creencias falsas deben descartarse. Las creencias deben ser cambiadas para ajustarse al mundo, no viceversa. Los deseos apuntan a ser realizados, y su realización consiste en que el mundo se ajuste a ellos. El hecho de que el contenido indicativo de un deseo no se realice en el mundo no es todavía un fracaso en el deseo, y no es ni siquiera una razón para descartar el deseo. El mundo, crudamente, debería ser cambiado para ajustarse a nuestros deseos y no a la inversa.

Yo creo que esta caracterización de la diferencia entre deseos y creencias captura algo profundo en nuestra concepción sobre su naturaleza. Más aún, quiero en definitiva argumentar que la idea de que los deseos son estados a los cuales debe el mundo ajustarse nos permite sacar a la luz una conexión importante entre nuestros conceptos de deseo y motivación.

Sin embargo, como advierte PLATTS, hablar de direcciones de ajuste de un estado es altamente metafórico. Esto es problemático. Podría pensarse que estamos injustificados en apelar a ese concepto al caracterizar a los deseos, y al iluminar la conexión entre deseos y motivación, si no tenemos un modo no metafórico de comprenderla. Más aún, como PLATTS advierte, si tomamos la caracterización en sentido estricto, no es claro si nos permite caracterizar a los deseos en absoluto. Pues, sostiene, dado que «todos los deseos parecen que involucrarán elementos de creencias», los deseos no son estados cuya dirección de ajuste sea enteramente del segundo tipo. Aparece entonces la pregunta acerca de si hay estados como esos (PLATTS, 1979: 257).

Me parece que PLATTS está en lo cierto al resaltar estos problemas con la metáfora, pero estaríamos equivocados al pensar que los problemas que él menciona son insuperables. Pues quiero sugerir que la caracterización metafórica de los deseos como estados de cosas tales que el mundo debe ajustarse a ellos se acomoda bien con otra sugerencia, más plausible, sobre la epistemología de los deseos, una sugerencia inspirada por otras observaciones de HUME.

HUME se daba cuenta perfectamente que de modo paralelo a las «pasiones violentas» que afectan al sujeto que las tiene, hay también «pasiones calmas», pasiones que carecen por completo de contenido fenomenológico. Estaba perfectamente al tanto de que su línea de argumento oficial sobre la epistemología de los deseos —que son conocidos por su fenomenología— era totalmente inadecuada como una teoría de la epistemología de las pasiones calmas y que por lo tanto necesitaba una teoría alternativa de la epistemología apropiada a esas pasiones. Como resultado, HUME sugirió que, por contraste con las pasiones violentas, las pasiones calmas «son más conocidas por sus efectos que por su sensación inmediata» (1888: 417).

Cuando STROUD considera esta sugerencia, señala que compromete a HUME a la idea de que los deseos deben ser concebidos como las causas de las acciones (STROUD, 1977: 166). Podría pensarse entonces que la sugerencia de HUME debería ser de poco interés para nosotros, pues he argumentado que el único argumento para la teoría humeana, si es que hay alguno, será que sólo ella es capaz de dar sentido a la explicación de razón como una especie de explicación teleológica, y que uno puede aceptar que las explicaciones de razón son teleológicas sin aceptar que las explicaciones de razón son causales. Pero si aceptamos esta concepción del deseo inmediatamente nos encerramos en una concepción causal.

Sin embargo, aunque esto hace que la aceptación de la sugerencia de HUME tal como está sea inapropiada, me parece a mí que haríamos mal en abandonar la sugerencia de HUME por completo. Pues si estamos menos interesados en la visión de HUME que en la visión humeana entonces parece

que podemos encontrar en la sugerencia de Hume sobre la epistemología apropiada para las pasiones calmas la inspiración para una concepción algo diferente de los deseos; una concepción que nos permita permanecer neutrales sobre la cuestión de si los deseos son causas.

De acuerdo con esta concepción alternativa, los deseos son estados que tienen cierto rol funcional. Esto es, de acuerdo con esta concepción, deberíamos considerar al deseo de hacer ϕ como el tener un conjunto de disposiciones, la disposición a hacer ψ en condiciones C, la disposición de hacer χ en condiciones C', etc., donde, para que tengan lugar las condiciones C y C', el sujeto tiene que tener, *inter alia*, algunos otros deseos, y por lo tanto algunas otras creencias medio-fin, creencias relativas a hacer ϕ haciendo ψ , hacer ϕ haciendo χ , etc. Pues la sugerencia de Hume sobre cómo las pasiones calmas son conocidas ahora puede ser traducida a la idea de que la epistemología del deseo es simplemente la epistemología de los estados disposicionales —esto es, la epistemología de esos contrafácticos. Esto no nos compromete, como Humeanos, a la tesis de que los deseos tienen que ser concebidos como las causas de las acciones. Pues es una tesis filosófica sustancial afirmar que las disposiciones son causas (PRIOR, PARGETTER y JACKSON, 1982)— aunque permítanme decir que mi propia visión de las disposiciones y por lo tanto de los deseos es que en realidad son causas (JACKSON y PETTIT, 1988).

Una concepción disposicional de los deseos nos permite resolver mucho de los problemas que hemos enfrentado hasta ahora. Por ejemplo, una concepción disposicional es precisamente la teoría de qué es un deseo que explica cómo puede ser que los deseos tengan contenido proposicional. Pues el contenido proposicional del deseo puede entonces ser simplemente determinado por su rol funcional (LOAR, 1981; JACKSON y PETTIT, 1988).

Una concepción disposicional de los deseos también satisface el requisito de una epistemología del deseo en favor del cual argumenté antes. Implica que la epistemología del deseo debe permitir que los sujetos sean falibles sobre los deseos que tienen. Pues, dada la suposición de que los deseos son disposiciones a actuar de cierto modo en ciertas condiciones, es implausible sugerir en líneas generales que, si los contrafácticos que son así verdaderos de un sujeto que desea hacer ϕ son verdaderos de él, entonces él cree que lo son, y es igualmente implausible sugerir generalmente que si un sujeto cree que esos contrafácticos son verdaderos de él, entonces esos contrafácticos son verdaderos de él⁶.

⁶ Así, yo sostengo que la creencia falsa de un sujeto de que desea hacer ϕ no es un estado que sea potencialmente explicativo de su comportamiento. La experiencia cotidiana avala esta postura. Reflexione sobre las ocasiones en que usted está parado en el borde de una piscina y piensa que tiene el deseo de zambullirse. En algunas de esas ocasiones, su cuerpo no responde en absoluto a los deseos que usted profesa tener. Mientras usted se queda allí inmóvil a veces llega

Más aún, una concepción disposicional de los deseos es consistente tanto con la afirmación de que ciertos deseos tienen contenido fenomenológico esencial como con la afirmación de que ciertos deseos carecen por completo de contenido fenomenológico. Pues, de acuerdo con esta concepción, los deseos tienen contenido fenomenológico justo en la medida en que tener ciertos sentimientos es una de las cosas que tienen una disposición a producir en ciertas condiciones. Algunos deseos pueden ser disposiciones a tener ciertas sensaciones en todas las condiciones: estos tienen contenido fenomenológico esencial. Otros deseos, aunque son disposiciones a comportarse en ciertos modos, pueden no ser disposiciones a tener ciertos deseos en absoluto: éstos carecen en absoluto de contenido fenomenológico.

También somos capaces, dado una concepción disposicional de los deseos, de ver por qué PLATTS está en lo cierto al afirmar que, de muchas maneras, los deseos «involucran elementos de creencia». Pues si el deseo de hacer ϕ es un cierto tipo de estado disposicional complejo del tipo descrito, entonces desear hacer ϕ puede «involucrar» elementos de creencia en cada uno de los siguientes modos: la obtención de las condiciones en las cuales el sujeto hace ϕ pueden requerir que tenga ciertas creencias, especialmente creencias medio-fin. La verdad del contrafáctico «si el sujeto estuviera en condiciones C haría ϕ » puede requerir que el sujeto tenga ciertas creencias dados requisitos holísticos sobre la atribución de deseo y creencia, etcétera.

Finalmente, una concepción disposicional nos permite ver por qué, a pesar del hecho de que los deseos pueden involucrar de muchas maneras elementos de creencia, podemos decir apropiadamente de alguien que tiene un deseo que está en un estado al cual debe el mundo ajustarse. Pues una concepción disposicional de los deseos nos permite presentar en términos no metafóricos, y por lo tanto a su vez encuentra apoyo de, la caracterización metafórica de las creencias y los deseos en términos de sus diferentes di-

a la conclusión de que, contrariamente a lo que pensó, no quería, en realidad, nadar después de todo. Así pues, tal como esperaríamos si esta afirmación es cierta, hay casos en que un sujeto cree que desea ϕ justo hasta el momento en que supone que actúa solo para descubrir que, de hecho, no tiene ese deseo cuando su cuerpo deja de responder a su deseo.

Por supuesto, podemos elaborar casos en los que puede parecer que las creencias falsas de un agente acerca de lo que quiere la motiven. El ejemplo de John el músico, tratado antes, quien seguía la carrera de músico a pesar de que no quería, a pesar de que falsamente creía que era su voluntad, podría ser un caso de éstos. Pero me parece que estamos obligados a pensar que en casos como éste la falsa creencia, en sí misma, no cumple ningún rol motivacional; que ese rol es cumplido por otros deseos que se encuentran presentes: por ejemplo, en este caso, el deseo de John de complacer a su madre. Estamos obligados a reconstruir los asuntos de este modo en pos de explicar la diferencia motivacional entre las dos maneras en que el caso de John podría haber terminado. Podría haber terminado tal como lo describí. O podría haber terminado como el caso del agente en la piscina. John podría haberse dado cuenta de que no era capaz de comprometerse con sus propósitos musicales. Su incapacidad podría haberle hecho reconocer que no quería ser un músico después de todo.

recciones de ajuste. Pues la diferencia entre deseos y creencias en términos de direcciones de ajuste puede ser vista como equivalente a la diferencia en los roles funcionales de la creencia y el deseo. Dicho de manera grosera, y simplificando un poco, equivale *inter alia*, a la diferencia en la dependencia contrafáctica de una creencia de que p y un deseo de que p de la percepción con el contenido de que no p : una creencia de que p tiende a desaparecer en la presencia de la percepción con el contenido de que no p , mientras que un deseo de que p tiende a persistir, disponiendo al sujeto en ese estado a lograr que p ⁷. Así, podríamos decir, las atribuciones de deseos y creencias requieren que diferentes tipos de contrafácticos sean verdaderos del sujeto a los cuales se atribuyen. Podríamos decir que en eso consiste la diferencia en sus direcciones de ajuste⁸.

⁷ ¿Por qué he dicho que esto es muy tosco y, de algún modo, simplificado? Porque donde « p » es una proposición sobre el pasado, o una proposición en el sentido de que algo sucede aunque no a través de la acción propia del agente, no será cierto que el deseo de p es una disposición para que se produzca p . Una caracterización más precisa y completamente general de rol funcional del deseo, necesitaría por ello ser dada a través de las líneas sugeridas en teoría de la decisión. Un deseo de que p es un estado que dispone a un sujeto a formular determinadas apuestas frente a la lotería, donde el resultado es *inter alia* que p . De acuerdo con esa teoría, es aún verdad que la diferencia entre el deseo y la creencia está dada por sus roles diferenciales en la explicación de la acción, aunque en este caso la acción consista en apostar. A pesar de tosca y simplificada, la sugerencia en el texto será suficiente para los propósitos presentes. Pues nuestro interés reside en explicar la naturaleza de las razones motivacionales, y los deseos que constituyen razones motivacionales son por ende deseos de estados de cosas que *no están* en el pasado, y que *han de ser* causados a través de la acción del sujeto. Tal como yo lo entiendo, esto constituye una respuesta a las críticas más insistentes de G. F. SCHUELER (1991) a una presentación más temprana de estas ideas (SMITH, 1987). Para una mayor argumentación de las direcciones de ajuste, véase HUMBERTONE (1992) y VELLEMAN (1992).

⁸ Philip PETTIT distingue entre dos tipos de estado que son tales que el mundo debe ajustarse a su contenido: *los deseos* y *los hábitos de inferencia* (1993: 18-19). Así, por ejemplo, él argumenta que el hábito de inferencia *modus ponens* es un estado que produce el resultado de creer que q cuando el *input* es una creencia de que p y una creencia de que $p \rightarrow q$, y, afirma él, esto es por tanto un estado con el que el mundo debe ajustarse tal como lo es un deseo; pues requiere que nuestra psicología sea de cierto modo, que nuestra psicología no sea así no es una razón para cambiar el hábito de inferencia, sino más bien una razón para cambiar nuestra psicología. PETTIT afirma que por eso necesitamos distinguir los deseos de los hábitos de inferencia y que las consideraciones de la dirección de ajuste son inadecuadas para esta tarea.

Su sugerencia positiva sobre cómo podríamos efectuar esta distinción es que los deseos, a diferencia de los hábitos de inferencia, son «canalizadas a través de creencias»: esto es, él sugiere que la creencia de que la respuesta es aceptable para realizar el contenido del hábito de inferencia no necesita canalizar el impacto del hábito de inferencia. Alguien que, deseando hacer ψ , hace ϕ , sólo hace ϕ porque su creencia de que puede hacer ψ al hacer ϕ canaliza el impacto de su deseo de hacer ψ . Pero alguien que tiene el hábito de inferencia *modus ponens* no necesita tener el impacto de su hábito canalizado de esta manera para terminar creyendo que q cuando cree que p y que $p \rightarrow q$. No hay necesidad de ningún rol jugado por la creencia de que creer q es una manera de creer algo que se sigue, *modus ponens*, de sus otras creencias.

Yo, sin embargo, no estoy convencido por la afirmación de PETTIT de que necesitamos distinguir los deseos de los hábitos de inferencia del modo que él sugiere. Lo que me parece que demuestra el ejemplo de los hábitos de inferencia es sólo cuántos cabos sueltos hay en la frase «dirección de ajuste», y que, para determinados propósitos, hagamos mejor en hablar directamente en términos de patrones de disposiciones, donde éstos logran explicarse en detalle en términos

Estos resultados son importantes. Pues permiten hacer que la concepción disposicional del deseo sea atractiva *simpliciter*, independientemente de la teoría de la motivación que favorezcamos. Más aún, nos permiten hablar sin vergüenza de los deseos como estados a los cuales el mundo debe ajustarse, pues esa forma de hablar, aunque metafórica, captura el rasgo que distingue los deseos de las creencias. Esta forma de hablar describe de modo apto el tipo de estado disposicional que un deseo es.

Las razones humeanas para creer P1 —el principio de que una razón motivacional está constituida por la presencia de un deseo y una creencia medio-fin— ahora pueden ser enunciadas de modo bastante simple. Dado que, como hemos visto, todos los teóricos deberían aceptar una concepción disposicional del deseo, y dado que esta concepción nos permite hablar de los deseos como estados a los cuales el mundo debe ajustarse, las razones humeanas son también, creo, tanto intuitivas como convincentes.

7. DESEOS, DIRECCIONES DE AJUSTE, OBJETIVOS Y RAZONES MOTIVACIONALES

¿Qué significa que alguien tenga una razón motivacional? El humeano dice que entendemos que significa que alguien tenga una razón motivacional en un momento al pensar en ese sujeto como, *inter alia*, teniendo un objetivo en ese momento (el *alia* aquí incluye tener una concepción de los medios para lograr el objetivo). Esto es, tener una razón motivacional es, *inter alia*, tener un objetivo. ¿Pero qué tipo de estado es tener un objetivo? ¿Qué dirección de ajuste tiene este estado? Ciertamente el tener un objetivo es un estado al que el mundo debe ajustarse, y no viceversa. Así, tener un objetivo es estar en un estado con la dirección de ajuste de un deseo, pero dado que todo lo que

funcionalistas. Pues parece suficientemente claro que, por ejemplo, el hábito de inferencia modus ponens debe ser él mismo interno a las disposiciones de creer que p , de creer que $p \rightarrow q$, y cosas por el mismo, y no, como lo parece sugerir el modo de ubicar las cosas de PETER, una disposición separada que se *combina con* las disposiciones constitutivas de estas creencias para producir otras creencias. Después de todo, ¿cuáles serían las disposiciones constitutivas de creer p si no incluyeran ya disposiciones como la disposición de creer q cuando se cree $p \rightarrow q$ y otras disposiciones similares? De hecho, parece suficientemente claro que los hábitos de inferencia deben ser concebidos de esta manera, porque ellos son en sí mismos criticables si no se ajustan al mundo. Alguien que tiene el hábito de inferencia de la falacia del jugador, por ejemplo, debe liberarse de ese hábito y adquirir en cambio un hábito diferente; un hábito que le permita mejor derivar verdades de las verdades. Y esto, a mi entender, es mejor entendido como una afirmación sobre las disposiciones constitutivas de ser un creyente. Alguien que tiene el hábito de inferencia de la falacia del jugador es un creyente defectuoso.

Cuando hablo de estados que tienen direcciones de ajuste, tengo entonces en mente paquetes enteros de disposiciones constitutivas de desear y de creer, no estados que pueden ser así descritos de modo más laxo. Por eso ignoro los hábitos de inferencia en lo que sigue.

puede decirse sobre un deseo es que es un estado con la dirección de ajuste apropiada, se sigue que tener un objetivo es simplemente desear⁹.

En resumen, entonces, el humeano cree que P1 porque P1 está implicado por las siguientes tres premisas:

- a) Tener una razón motivacional es, *inter alia*, tener un objetivo.
- b) Tener un objetivo es estar en un estado al que el mundo debe ajustarse.
- c) Estar en un estado al que el mundo debe ajustarse es desear.

Aunque este argumento es simple, creo que es realmente poderoso. Después de todo, ¿qué premisa en el argumento podría ser plausiblemente negada? Considerémoslo en orden.

Dado el supuesto de que las explicaciones de razón son explicaciones teleológicas a) parece invulnerable; de hecho tiene el estatus de una verdad conceptual, pues entendemos qué significa que alguien tiene una razón motivacional en parte precisamente porque la consideramos como teniendo un objetivo. Y b) es invulnerable también, por razones similares. Pues darse cuenta del hecho de que el mundo no es como está especificado en el contenido de nuestro objetivo es suficiente para no abandonar ese objetivo, es suficiente más bien para cambiar el mundo. La premisa más vulnerable es quizás c), la afirmación de que estar en un estado al que el mundo debe ajustarse es desear. Puedo imaginar dos tipos de objeción a esta premisa.

Primero, de acuerdo a PLATTS, ANSCOMBE afirma sólo que el deseo es «un ejemplo central» de aquellos estados a los que el mundo debe ajustarse. Pero, podría decir, hay otros estados que tienen esa dirección de ajuste también: las esperanzas, las intenciones, el querer, etc. Por lo tanto, dado a) y b), con seguridad debemos decir que tales estados pueden constituir el tener objetivos también. Pero, *ex hypothesi*, las esperanzas, el querer y las intenciones no son deseos.

Sin embargo, un ataque al argumento humeano de este tipo claramente no es un ataque al *espíritu* del argumento, es más bien un ataque sobre los

⁹ ¿Y qué tal tener a la verdad como meta? ¿Implica esto el desear la verdad? Y si es así, dado que los creyentes tienen a la verdad como meta, ¿se sigue de ahí que los creyentes deben ser «deseantes» de la verdad? Yo no lo pienso así. Tal como lo veo, hablar de que los creyentes tienen la verdad como meta se entiende mejor de modo más laxo. Decir que un creyente tiene la verdad como meta no es asignarle disposiciones *adicionales* a las disposiciones constitutivas de ser un creyente. Más bien, las disposiciones constitutivas de tener a la verdad como son en sí mismas *internas* a las disposiciones constitutivas de ser un creyente. Así, según lo veo yo, es un error suponer que haya un creyente que no tenga a la verdad como meta. Alguien que no tiene a la verdad como meta no es un *creyente* legítimo en absoluto. Ése es el verdadero alcance de la afirmación de que las creencias son estados que deben ajustarse al mundo. Véase nuevamente la nota 8 para una discusión relevante.

detalles del argumento. El humeano podría, por supuesto, optar por defender los detalles. Podría argumentar que, aunque estos estados en algún sentido tienen la dirección de ajuste de un deseo, en la medida en que difieren de los deseos, difieren porque también involucran elementos de creencia. Así podría sugerir que tales estados son de hecho compuestos de deseo y creencia: querer que p es simplemente desear que p y creer que p es muy improbable. Tener la intención de hacer φ es simplemente desear intensamente hacer φ y creer que por lo tanto intentará hacer φ ; etc. Alternativamente, el humeano podría aceptar que, aun después de quitar los elementos de creencia en tales estados, es implausible suponer que el estado que queda es él mismo un deseo. Podría por lo tanto elegir conceder los detalles del argumento al objetor. Pues si «deseo» no es una categoría lo suficientemente amplia del estado mental que abarque todos aquellos estados con la dirección de ajuste apropiada, entonces el humeano puede simplemente definir el término «pro-actitud» como «el estado psicológico al cual el mundo debe ajustarse». Y después afirmar que las razones motivacionales están constituidas, *inter alia*, por pro-actitudes (compárese DAVIDSON, 1963: 4).

Una segunda objeción a *c*) concede que las creencias y los deseos tienen sólo una dirección de ajuste, pero ataca más bien el supuesto de que sólo hay estados que tienen una o la otra dirección de ajuste. El supuesto de que podemos siempre quitar los elementos similares-a-creencia de los estados que tienen la dirección de ajuste de un deseo¹⁰. El objetor quiere saber por qué no podría haber un estado de algún otro tipo que tenga ambas direcciones de ajuste; un estado que es a la vez tal que el mundo debe ajustarse a él y que debe ajustarse al mundo. Si tal estado fuera posible entonces, podría decir, constituiría tener un objetivo. Pero dado que no sería un deseo, y ni siquiera una pro-actitud, y ya que se nos ha concedido que los deseos y las pro-actitudes tienen sólo una dirección de ajuste, deberíamos entonces rechazar *c*). Las razones motivacionales podrían estar constituidas por este tipo adicional de estado.

Sin embargo, aunque pueda sonar como una posibilidad coherente que haya tal estado, no lo es en realidad, al menos no si tomamos la sugerencia

¹⁰ ¿Debe el antihumeano conceder que las creencias y los deseos tienen solamente una dirección de ajuste? Como argumenté en otras oportunidades, en un sentido esta cuestión es enteramente terminológica (PETTIT, 1987; SMITH, 1988a). Sin embargo, para responder la pregunta discutida más adelante en el texto, «¿hay estados psicológicos que son similares-a-creencias y similares-a-deseos?», debemos tomar una posición sobre la cuestión terminológica. Yo interpreto esto como una pregunta acerca de si hay algo como un tercer tipo de estado psicológico —*creseos* (ALTHAM, 1986) o cuasi-creencias tal como las denominé en otras oportunidades (1986:56)— no como una pregunta acerca de si algunas creencias son también deseos. Nótese, sin embargo, que si elegimos formular esta pregunta en estos últimos términos, entonces deberíamos reformular el principio humeano P1 como una afirmación de que las razones motivacionales están constituidas, *inter alia*, por deseos-que-no-son-creencias (SMITH, 1988a).

de modo bastante literal. Pues, tal como hemos entendido el concepto de dirección de ajuste, la dirección de ajuste de un estado con el contenido de que p está determinada, *inter alia*, por su dependencia contrafáctica de una percepción con el contenido de que no p . Un estado con ambas direcciones de ajuste debería entonces ser tal que, en la presencia de tal percepción tiende a desaparecer, y a la vez, en la presencia de esa percepción tiende a persistir, llevando al agente que lo tiene a lograr que p . Tomada de modo bastante literal, entonces, la idea de que puede haber un estado que tiene ambas direcciones de ajuste es completamente incoherente.

Un objetor más sutil podría encontrar defectuosa esta réplica (McDOWELL, 1978; McNAUGHTON, 1988). Podría decir que la réplica funciona sólo si consideramos que la sugerencia es que hay un estado que tiene las dos direcciones de ajuste con respecto al mismo contenido. Pero, se podría preguntar, ¿por qué tomar la sugerencia literalmente? Después de todo, ¿por qué no podríamos pensar que hay, además de los estados que tienen una u otra dirección de ajuste respecto de un mismo contenido, estados unitarios que tienen ambas direcciones de ajuste, aunque respecto de dos contenidos diferentes? El objetor incluso podría ofrecernos un ejemplo de tal estado —de modo no sorprendente, un ejemplo que nos lleva de vuelta al problema moral original, y a su solución preferida a dicho problema—.

Considérese qué deberíamos pensar de modo ordinario sobre la creencia moral de que es correcto hacer φ . El objetor podría decir que esto es realmente mejor visto como, para usar el término excelente de J. E. J. ALTHAM, el *creseo* de que hacer φ es correcto —«*creseo*» porque, como veremos aunque este estado es similar a la creencia, también es similar al deseo (1986)—. El *creseo* de que φ es correcto es apropiadamente descrito como el estado que debe ajustarse al mundo porque tiende a desaparecer cuando el sujeto está confrontado con la percepción del contenido de que hacer φ es incorrecto. Ésta es la razón por la que, en términos del problema moral original, estábamos dispuestos incorrecta, aunque comprensiblemente, a decir que los juicios morales expresan creencias. Más aún, el objetor podría señalar, el *creseo* de que hacer φ es correcto también es apropiadamente descrito como un estado al que el mundo debe ajustarse, pues que un sujeto tenga el *creseo* de que hacer φ es correcto lo dispone a hacer φ . Los *creseos*, un tercer tipo unitario de estado psicológico además de la creencia y el deseo, son *tanto* expresados en nuestros juicios morales del tipo «hacer φ es correcto» como constitutivos de nuestras motivaciones morales. ¿Cómo habremos de replicar a este objetor¹¹ más sutil?

¹¹ David LEWIS sostiene que el concepto de un *creseo* instrumental es incoherente por inconsistente con los postulados de la teoría de la decisión bayesiana (1988). Una versión del argumento de LEWIS para la teoría de la decisión no cuantitativa es desarrollada por John COLLINS (1988).

Seamos claros sobre qué está en juego. Los humeanos no necesitan negar que los agentes pueden, por ejemplo, creer que es correcto hacer ϕ y desear hacer ϕ : esto es, no necesitan negar la coexistencia contingente de deseos y creencias. Y tampoco, como veremos en el próximo capítulo, necesitan negar que si un agente cree que es correcto para él hacer ϕ , entonces debería racionalmente hacer ϕ : esto es, no necesitan negar que la coexistencia contingente de ciertas creencias y deseos es racionalmente requerida (SMITH, 1988a, 1988b). Más bien, lo que los humeanos deben negar y niegan es simplemente que los agentes que están en estados similares-a-las-creencias y estados similares-a-los-deseos están en un tipo de estado singular y unitario. Éste es el valor efectivo de la doctrina humeana según la cual el deseo y la creencia son existencias diferentes. Y su argumento para esta afirmación es realmente bastante simple. Es que es siempre al menos posible para los agentes que están en un estado similar-a-la-creencia no estar en algún estado particular similar-al-deseo, que los dos pueden siempre separarse, al menos modalmente. Esto, de acuerdo con los humeanos, es la *razón* por la cual son existencias diferentes.

En el caso particular bajo discusión, entonces, la afirmación humeana debe ser que es siempre al menos *posible*, para los agentes que están en un estado similar-a-la-creencia de que hacer ϕ es correcto, que a pesar de ello no tengan un estado similar-al-deseo de hacer ϕ . Que los dos pueden ser siempre separados, al menos modalmente. Y, correspondientemente, los antihumanos deben decir precisamente lo opuesto. Deben afirmar que es *imposible* para los agentes que están en un estado similar-a-la-creencia de que hacer ϕ es correcto no estar en un estado similar-al-deseo de hacer ϕ . Que los dos no pueden separarse, ni siquiera modalmente. Como yo lo veo, el desacuerdo entre humeanos y antihumanos acerca de si somos sólo agentes que creen y que desean, o también que *cresean*, equivale ni más ni menos que a un desacuerdo sobre estas afirmaciones modales.

Sin embargo, una vez que la cuestión es planteada en estos términos se vuelve claro que el antihumano está en un terreno inestable. Todo depende

No obstante, incluso si LEWIS y COLLINS estuviesen en lo cierto sobre esto [véase la respuesta de HUW PRICE (1989)], no sería suficiente para demostrar que no hay *creseos* en absoluto. Pues para argumentar en pos de esa conclusión necesitamos un argumento adicional, un argumento en el sentido de que tampoco hay *creseos* no instrumentales. Lewis sí intenta brindar tal argumento, pero no es convincente (1989: nota 1). Los *creseos* no instrumentales son, según nos dice él, «inmunes a cambiar bajo el impacto de la experiencia». Pero por qué él piensa esto no lo sé. Seguramente, si hubiese *creseos* no instrumentales, todo lo que se seguiría es que ellos cambian bajo el impacto de la experiencia del mismo modo en que lo hace una creencia ordinaria, no un deseo instrumental. Así, las experiencias que son suficientes para deshacerse de la parte similar-a-creencia también serán suficientes para deshacerse de la parte similar-a-deseo; y las experiencias que son suficientes para adquirir la parte similar-a-creencia serán suficientes también para adquirir la parte similar-a-deseo.

de cómo deberíamos interpretar la idea de que el juicio moral es esencialmente práctico. ¿Acaso la idea es que alguien que juzga correcto actuar de cierto modo está motivado a actuar de modo conforme *simpliciter*? ¿O es más bien la idea que alguien que juzga que actuar de cierto modo es correcto está motivado a actuar de modo conforme *ceteris paribus*? El antihumano necesita que el juicio moral sea práctico en el primer sentido. Pero nuestro juicio moral es de modo bastante evidente práctico sólo en el segundo sentido. (Recuérdese de nuevo las dos primeras versiones del internalismo descriptos al principio del capítulo II). Contrariamente a la primera idea, las irracionalidades prácticas de distintos tipos —debilidad de la voluntad y cosas por el estilo— pueden romper la conexión entre juicio moral y motivación.

Este punto es bien expresado por Michael STOCKER en su artículo «Desiring the Bad: An Essay in Moral Psychology».

A través del cansancio espiritual o físico, a través de la pereza espiritual, la debilidad del cuerpo, la enfermedad, la apatía general, el desaliento, la inhabilidad para concentrarse, un sentimiento de inutilidad o futilidad y cosas por el estilo, uno puede sentirse menos y menos motivado a buscar lo que es bueno. El deseo atenuado de uno no necesita ser una señal, y mucho menos ser el producto de, el hecho de que, o la creencia de que, hay menos bien a ser obtenido o producido, como el caso del Weltschmerz universal. De hecho, un defecto frecuente añadido de estar en tales «depresiones» es que uno ve todo el bien que podría ganarse o conservarse y a uno le falta la voluntad, el interés, el deseo o la fuerza (STOCKER, 1979: 744).

Y lo mismo vale para la conexión entre creer que una acción es correcta y estar motivado a llevarla a cabo. Es un lugar común, un hecho de la experiencia moral ordinaria, que las irracionalidades prácticas de distintos tipos —varios tipos de «depresión» como dice STOCKER— pueden dejar a nuestra visión evaluativa intacta y al mismo tiempo remover por completo sus motivaciones. La afirmación de los antihumanos de que los juicios morales son expresiones de *creseos* es inconsistente con esto. La afirmación humeana según la cual son expresiones de creencias no lo es. La opinión de los antihumanos debe por lo tanto ser rechazada en favor de la de los humeanos.

En este punto el antihumano insistirá que afrontemos directamente la teoría de John McDOWELL de la psicología de la persona virtuosa. Pues él insistió célebramente en que la gente virtuosa tiene un modo distintivo de concebir las situaciones que enfrenta. Concebir las situaciones de ese modo, nos dice McDOWELL, es estar en un estado cognitivo cuya «posesión implica una disposición de la voluntad del poseedor» (1978: 18). En otras palabras es tener un *creseo*. Él argumenta en favor de esta afirmación a través de un ejemplo.

[Al] urgir una conducta que uno considera moralmente requerida, uno se encuentra diciendo cosas como esta: «Usted no sabe lo que significa que alguien sea tímido y sensible». Transmitir lo que una circunstancia significa en este sentido cargado, es llevar a alguien a ver las cosas del modo en que una persona virtuosa las vería. En el intento por hacerlo, uno explota ardidés similares a aquellos que uno explota en otras áreas en que las que la tarea es reforzar la exhortación «véalo así»: yuxtaposiciones útiles de casos, descripciones con términos cuidadosamente elegidos y énfasis cuidadosamente puestos, y cosas por el estilo [...] ninguno de esos ardidés puede asegurar el éxito, en el sentido de que el fracaso mostraría irracionalidad de parte de la audiencia. Eso, junto con la importancia del uso exitoso de habilidades retóricas, los distingue del tipo de cosas que típicamente consideramos paradigmas de argumentos. Pero esto parece fundamento insuficiente para concluir que son apelaciones a la pasión como opuestas a la razón: para concluir que «véalo de este modo» es realmente una invitación encubierta a sentir, más allá de nuestra visión de los hechos, un deseo que se combinará con la creencia de uno para recomendar actuar de modo apropiado. (1978: 21-22).

Ahora bien, McDOWELL en realidad hace dos cosas aquí. Permítaseme separarlas. Primero, McDOWELL describe, en lo que me parecen términos más o menos precisos, un proceso ordinario de argumento moral. Cuando nos involucramos en tal argumento nos involucramos, tal como él sugiere, en varias formas de razonamiento analógico. En este caso particular, intentamos que nuestra audiencia llegue a pensar que alguien sensible y tímido es de modo relevante similar a otros modos en que la gente podría ser, modo que nuestra audiencia considera que justifican un trato más especial o más sensible. Este razonamiento yace en el centro mismo de los procedimientos a través de los cuales intentamos lograr conocimiento moral, yace en el centro mismo de lo que antes describimos bajo el rótulo de «equilibrio reflexivo» (capítulo II).

Segundo, sin embargo, y de modo más controvertido, McDOWELL nos ofrece una *interpretación* de qué hacemos cuando nos involucramos en este tipo de razonamientos. Específicamente, él sugiere que estamos tratando de lograr que nuestra audiencia tenga un *creseo*. La verdadera cuestión es por lo tanto si su discusión del ejemplo hace que esa interpretación sea plausible. En otras palabras, ¿nos provee con un ejemplo plausible de un estado cognitivo cuya posesión implica una disposición de la voluntad del poseedor: la concepción de la persona virtuosa de alguien tímido y sensible? No lo creo.

McDOWELL ha sido llevado a confusión, creo, por el hecho de que, si aceptamos las suposiciones morales sustantivas que él formula, entonces aceptaremos que un condicional de este tipo es verdadero.

(x) (Si x es virtuoso y x concibe a alguien como tímido y sensible, entonces x está dispuesto a hacer que esa persona se sienta cómoda en su compañía, a protegerla de aquellos que la incomodan, etc.).

La lista por supuesto será abierta porque puede no haber límite a las diferentes cosas que deberían hacerse en diferentes circunstancias. Ahora bien, yo acuerdo que aceptaremos este condicional, ¿pero acaso aceptarlo requiere aceptar también que la persona virtuosa tiene un modo especial de concebir las situaciones que enfrenta? ¿Equivale a aceptar que estar en un estado similar-a-la-creencia de que alguien es tímido y sensible es estar en un estado similar-al-deseo de que debemos hacerlo sentir confortable en nuestra compañía, protegerlo de aquellos que lo incomodan, etc.? ¿Es eso lo que el condicional dice? No. No es eso lo que el condicional dice. Lo que el condicional dice es que hay una conexión necesaria en el caso restringido en que la persona que cree es una persona virtuosa. Y si, tal como dirá el humeano, alguien cuenta como una persona virtuosa sólo si tiene ciertos deseos —el deseo de hacer que alguien que es tímido y sensible se sienta confortable en su compañía, el deseo de protegerlo cuando alguien lo incomoda, etc.— entonces vemos inmediatamente que la conexión necesaria es un artefacto de esa restricción. La persona virtuosa es una persona que cree y desea de modo regular después de todo.

De hecho, me parece que el humeano puede dar un argumento concluyente en favor de esta afirmación (WALLACE, 1983). Pues si la apreciación de la persona virtuosa del hecho de que alguien es tímido y sensible fuera realmente un *creseo*, debería ser imposible separar la parte similar-a-la-creencia de la parte similar-al-deseo; debería ser imposible, por ejemplo, que la persona virtuosa retenga su apreciación del hecho de que alguien es sensible y tímido y sin embargo no esté en un estado similar-al-deseo de hacer que esa persona se sienta confortable en su compañía. Esto es, en efecto, insistir en que aquellos que alguna vez fueron virtuosos, que a través de algún evento desafortunado u otro, sufren de debilidad de la voluntad en alguna ocasión, no deben apreciar ya más el hecho de que alguien es tímido y sensible; de que tienen que haber olvidado algo que solían conocer muy bien (McDOWELL, 1979). Pero esto es seguramente bastante poco creíble. Las observaciones de STOCKER sobre los efectos de la «depresión» son otra vez apropiadas aquí. Es un lugar común, un hecho de la experiencia moral ordinaria, que cuando los agentes sufren de debilidad de la voluntad pueden mirar de frente a los hechos que solían motivarlos, apreciarlos en toda su gloria, y sin embargo no estar motivados por ellos (PETTIT y SMITH, 1993a). Que esto sea así es crucial para una comprensión de cuán horrible puede ser la debilidad para algunas personas. La descripción de McDOWELL de la persona virtuosa es simplemente inconsistente con este lugar común. Si quiere

tener éxito en su argumento, tiene por lo tanto que convencernos que hay algún hecho u otro que no puede simplemente ser visto directamente de este modo por alguien que sufre debilidad de la voluntad.

McDOWELL de hecho enfrenta esta dificultad. Él admite que, por cualquier examen ordinario, las personas virtuosas y no virtuosas conciben las mismas proposiciones.

El fracaso en ver qué significa una circunstancia, en el sentido cargado, es por supuesto compatible con ser competente, de acuerdo a todos los *test* ordinarios, en el uso del lenguaje empleado para describir la circunstancia; esto hace ver cuán cargada está la noción de significado involucrado en la protesta (1978: 22).

Pero él piensa que esto simplemente muestra que los *test* ordinarios que tenemos para identificar las proposiciones afirmadas por los agentes son inadecuados.

Para preservar la distinción deberíamos decir que las concepciones relevantes *no son algo así como poseídas excepto por aquellos cuyas voluntades son influenciadas apropiadamente* (1978: 23, la cursiva es mía).

En otras palabras, él piensa que podemos simplemente definir las proposiciones afirmadas por la persona virtuosa como las proposiciones que son tales que, si alguien está en un estado similar-a-la-creencia respecto de dicha proposición, entonces está en un estado similar-al-deseo con respecto a otra proposición correspondiente.

Pero si bien esta maniobra está técnicamente disponible, sólo sirve para dejar en claro cuán intratable es el problema que en realidad McDOWELL enfrenta. Pues lo que ya hemos visto es que todos los *test* ordinarios a través de los cuales identificamos proposiciones, y los hechos de nuestra experiencia moral ordinaria —esto es, las descripciones de STOCKER de las varias «depressiones» que socavan la motivación pero dejan la cognición intacta— sugieren que no hay tales proposiciones. Pero en ese caso es difícil ver qué razón posible podría tener McDOWELL para pensar que hay tales proposiciones. Irónicamente, parece que es la visión antihumana de McDOWELL la que se ha tornado en un dogma, no la visión humeana que supuestamente iba a reemplazar.

Éste es un resultado agradable. Pues es porque él piensa que hay ciertas proposiciones que son tales que, si son en realidad afirmadas, entonces quien las afirma debe estar dispuesto a actuar de un cierto modo, que McDOWELL piensa que puede conceder P2 al humeano —la afirmación de que la motivación requiere la presencia de un deseo— y a la vez rechazar P1 —la afirmación de que estos deseos constituyen nuestras motivaciones—. Su idea es ésta. Ya que, si un agente se encuentra en un estado de *creseo*, es aceptable

adscribirle el correspondiente deseo también, así podemos conceder que la motivación requiere la presencia de un deseo, y entonces P2 (1978: 25). Pero en la medida en que recordemos que el deseo así adscrito es meramente «adscrito consecuencialmente» —esto es, que su adscripción aceptable es una mera consecuencia del hecho de que el agente tiene un *creseo*— vemos que no estamos por ello comprometidos a pensar que el deseo del agente constituye su motivación: esto es, no estamos comprometidos con P1. Más bien podríamos decir que su motivación está constituida por su *creseo*.

Sin embargo, si el argumento que hemos dado hasta aquí es correcto, sale a la luz que P2 no es un principio en relación al cual debemos tener ninguna razón para creer por derecho propio. No hay proposiciones del tipo imaginado por McDOWELL. Entendido apropiadamente, entonces, nuestra única razón para sostener P2 es que está implicado por P1. Y, como hemos visto, tenemos todas las razones para creer en P1. Los deseos son efectivamente los únicos estados que pueden constituir nuestras razones motivacionales¹².

8. RESUMEN DEL ARGUMENTO HASTA AQUÍ Y ANTICIPO

En este capítulo he argumentado en favor de la teoría humeana de la motivación. Esta teoría comienza, espero, a parecer mucho más plausible de lo que muchos han pensado. El argumento que he dado ha sido en realidad bastante simple. Una vez que vemos con firmeza que la teoría humeana es una teoría de las razones motivacionales, y nos equipamos con una concepción adecuada del deseo, vemos que sólo la afirmación humeana de que los deseos son constituidos, *inter alia*, por deseos, es capaz de dar sentido apropiadamente a la explicación de razón como una especie de explicación teleológica.

A pesar de que este argumento es simple parece que ha sido completamente pasado por alto por aquellos teóricos antihumanos de la motivación

¹² Como yo lo entiendo, la visión de NAGEL es similar a la de McDOWELL. NAGEL piensa que determinadas proposiciones con base impersonal e intemporales son tales que, si son aceptadas en absoluto, entonces quien las acepta está dispuesto a actuar de determinada manera. Así, por ejemplo, NAGEL piensa que si ahora yo acepto la proposición de que A está dolorido en el tiempo *t* —lo acepto y al mismo tiempo aprecio completamente la realidad de otras personas y tiempos— entonces yo estoy ahora dispuesto a aliviar el dolor de A por *t*. Sin embargo, NAGEL también sostiene que simplemente en virtud de estar así dispuesto podemos decir que el agente tiene un «deseo motivado» para perseguir el acto en cuestión. Por esta razón, él piensa que deberíamos conceder P2 a los humanos: la motivación efectivamente conlleva la presencia del deseo. Pero, como él apunta, «si el deseo es un deseo motivado, su explicación será la misma que la explicación de la búsqueda, y no es para nada obvio que un deseo deba incluirse en esta explicación ulterior» (1970: 29). La búsqueda podría explicarse mejor apropiadamente por la presencia de un *creseo* y no de un deseo. Y, de ser así, deberíamos negar P1. Las motivaciones pueden estar constituidas por *creseos*; no solamente están constituidas por deseos. Así lo dice NAGEL. Y nuestra respuesta a NAGEL es la misma que la hecha a McDOWELL en el texto.