



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN FILOSOFÍA
FACULTAD DE FILOSOFÍA Y LETRAS

**«La flexibilidad de los sentimientos»:
La teoría humeana de la motivación y de los juicios morales**

TESIS
QUE PARA OPTAR POR EL GRADO DE
DOCTOR EN FILOSOFÍA

PRESENTA:
FRANCISCO JAVIER SERRANO FRANCO

DIRECTORA DE TESIS
DRA. FAVIOLA RIVERA CASTRO
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS DE LA UNAM

MIEMBROS DEL COMITÉ TUTORAL
DRA. MARGARITA VALDÉS VILLAREAL
DR. GUSTAVO ORTIZ MILLÁN
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS DE LA UNAM

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO, FEBRERO DE 2017



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A Rocío, la fuente de mis mejores pasiones

Estoy muy agradecido con la Dra. Margarita Valdés Villareal, el Dr. Gustavo Ortiz Millán y la Dra. Faviola Rivera Castro, estupendos maestros y amigos

Índice

Obras de Hume citadas y abreviaturas	6
Introducción.....	8
1. Las determinaciones naturales que explican a las creencias sobre hechos	16
1.1 Las determinaciones básicas de la mente.....	17
1.2 Las impresiones sensibles y las creencias sobre hechos derivadas directamente de ellas	23
1.3 Las creencias sobre hechos y la «definición» de creencia	24
1.4. La influencia o función <i>práctica</i> de la creencia	29
1.5 Las creencias causales y el papel de la razón en la formación de creencias	34
1.5.1 El punto de vista negativo: la razón no interviene en la formación de creencias causales	39
1.5.2 El punto de vista positivo: la costumbre, o la experiencia, explican a las creencias causales.....	42
1.5.3 La discusión de la probabilidad como origen de las creencias causales	45
2. Elementos causales y cognitivos de la teoría de las pasiones de Hume.....	51
2.1. Las pasiones calmadas y el deseo de belleza	52
2.2. Las ideas, las sensaciones y las determinaciones de las pasiones directas e indirectas	54
2.2.1 El contenido representacional de las pasiones indirectas.....	66
2.2.2. La interpretación de Alanen de la teoría de las pasiones de Hume	73
2.2.3. Simpatía e imaginación en la teoría de las pasiones de Hume	75
3. La teoría de la motivación de Hume	79
3.1. Razones, pasiones y deseos en la formación de motivos para actuar.....	81

3.2. La teoría instrumentalista de la motivación de Hume.....	91
3.3. La teoría de la motivación de Hume y la acusación de escepticismo de Korsgaard.....	94
3.4. La respuesta al reto de la “verdadera irracionalidad” de Korsgaard: una teoría humeana de la <i>deliberación</i> práctica. Las ideas de Bernard Williams al respecto.....	97
3.5. El argumento de Smith para defender una interpretación humeana de la motivación: «el observador ideal»	111
4. La teoría de los juicios morales de Hume	122
4.1. Argumentos con los que Hume prueba que las distinciones morales no se derivan de la razón.....	122
4.2. La interpretación <i>sentimentalista</i> del sentimiento moral	128
4.3. La interpretación del <i>observador ideal</i> del sentimiento moral.....	132
5. El naturalismo moral de Hume: la agencia y la moral	148
5.1. La moral y la psicología individual.....	149
5.2. Cómo los «artificios» o convenciones refuerzan a las determinaciones morales naturales	156
5.3. Cómo los agentes conforman su conducta a las reglas morales.....	163
5.4. Cómo las respuestas de los agentes morales son consistentes entre sí: los elementos normativos de la teoría moral de Hume	174
6. Consideraciones finales: Valoración del naturalismo moral de Hume y el papel que da a la belleza y a la utilidad. Defensa de la interpretación del observador ideal	183
6.1. La utilidad, las reglas morales y el sentimiento moral	187
6.2. La utilidad, la amabilidad y la belleza en el naturalismo de Hume.....	190
6.3. La teoría del observador ideal de Hume y las prácticas morales	195
7. Bibliografía.....	201

Obras de Hume citadas y abreviaturas

- Hume, David (1978), *A Treatise of Human Nature*, ed. por Selby-Bigge. Oxford: Clarendon Press (2a ed). [T]
- Hume, David (2000), *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton, Oxford/New York: Oxford University Press, 2000. [T]
- Hume, David (1988), *Tratado de la naturaleza humana* (Tr. y ed. de F. Duque). Madrid: ed. Tecnos. [T]
- Hume, David (1999), *An Enquiry Concerning Human Understanding*, ed. por T. L. Beauchamp. Oxford: OUP. (Col. Oxford Philosophical Texts), [EEH]
- Hume, David (1988), *An Enquiry Concerning the Principles of Morals*, ed. por T. L. Beauchamp, Oxford: Oxford University Press. [IPM]
- Hume, David (2008), *Ensayos morales y literarios* (Ed. y tr. de E. Trincado). Madrid: ed. Tecnos.
- Hume, David (2011), *Ensayos morales, políticos y literarios* (Ed. de E. Miller. Tr. de C. M. Ramírez. Madrid: Ed. Trotta.
- Hume, David, “De la norma del gusto” en Hume (2011), pp. 219-237. [DNG]

Puede que no esté ahora de más el poner de relieve la flexibilidad de nuestros sentimientos y las distintas modificaciones que tan fácilmente reciben de los objetos con que están conectados.

David Hume, *Tratado de la naturaleza humana* 3.3.6.6 SB 617

Introducción

Con este trabajo sobre el pensamiento de David Hume me concentro en sus ideas sobre la motivación y la moralidad. Al respecto, pretendo extender la llamada interpretación del «observador ideal» para cubrir aspectos de su pensamiento moral que no suelen tratarse bajo ella. Los defensores del observador ideal, según Elizabeth Radcliffe, sostienen que para Hume las distinciones morales están basadas en los sentimientos hipotéticos de un espectador ideal, alguien completamente informado, enteramente objetivo y no auto-interesado» (cf. Radcliffe 1994, 37). Esta posición se opone a la interpretación sentimentalista, según la cual las distinciones morales son, para Hume, meramente la expresión de ciertos sentimientos de aprobación y desaprobación (v. abajo capítulo 4). Usualmente, estas interpretaciones se refieren a los juicios morales o distinciones morales. Con «el observador ideal» se pretende resolver varias dificultades de la interpretación sentimentalista, por ejemplo, la aparente falta de criterios normativos o de deseabilidad de los fines de las acciones de los agentes. En efecto, si los juicios morales se refieren a los fines de ciertas acciones y si son meramente la expresión de ciertos sentimientos, ¿cómo podría un agente dado evaluar tales sentimientos o fines? Un rasgo de los sentimientos es que, siguiendo una caracterización de Hume, son «existencias originales», esto es, modificaciones de la consciencia que no son en absoluto objeto de consideraciones racionales. Los agentes son pasivos ante los sentimientos. Éstos les sobrevienen y no pueden hacer nada para modificar su manifestación en la consciencia. De manera que si un juicio moral es la expresión de una sensación placentera (positiva), o dolorosa (negativa), un agente no podría modificar este sentimiento y su conducta se vería dirigida sin remedio por él, aun cuando podrían ofrecerse consideraciones disuasorias de peso. Otra importante dificultad para la interpretación sentimentalista se encuentra en el terreno de la meta-ética.

En efecto, si los juicios morales son expresión de sentimientos, entonces tienen que ser sensibles a ciertos objetos externos al agente que los experimenta. Un sentimiento es una reacción a cierto objeto, como el frío o una amenaza (cuando además se combina con una pasión, como en el caso del miedo). Así, cabe preguntar: ¿a qué se refieren los sentimientos morales? ¿A qué propiedades o características? ¿Qué son esas supuestas propiedades o características, cuál es su naturaleza...? La interpretación sentimentalista enfrenta esto como una dificultad porque la idea

de que los juicios morales son expresiones de sentimientos morales implica que son manifestaciones *subjetivas* de las apreciaciones morales de los agentes y, al mismo tiempo, esta formulación parece implicar que tales sentimientos son *acerca de algo*, como en el caso de otros sentimientos. De modo que es posible que varios agentes coincidan en la misma apreciación si ese algo es una propiedad moral distintiva de ciertos objetos, como podrían ser el carácter de alguna persona o ciertas acciones percibidas como generosas, bondadosas o buenas. Además, la segunda implicación es compatible con la idea de que tales acciones tengan aquellas propiedades morales necesarias para ser percibidas por los agentes *independientemente* de quien sea quien las realice, etc. Así, el subjetivismo de los juicios morales parece que abre paso a un cierto *objetivismo* o a variables como el *proyectivismo*, esto es, en la medida en que pueda establecerse que las propiedades a las que se refieren los sentimientos sean, en efecto, objetivos o que se relacionen apropiadamente con ciertas propiedades intrínsecas a ciertos objetos. Discuto estas dificultades con detalle en la sección 4.2. Por ahora, sostengo que la interpretación del observador ideal promete responder a ellas, al tiempo que mantiene la idea de que los juicios morales son expresión de sentimientos.

El «observador ideal» es, realidad, un espacio psicológico donde el agente delibera sobre qué siente, qué debería hacer o qué debería valorar. Sin embargo, este proceso no es racional y en su ejecución el agente no sigue pautas de pensamiento deductivo. La posición que defiendo en esta investigación es que éste es un proceso muy amplio que se refiere no sólo a los juicios y valoraciones de un agente, sino, que abarca prácticamente todo lo que piensa, siente, cree y hace (v. abajo sección 6.1 y D. Wiggins 2006, p. 35). En este sentido digo que mi interpretación cubre más elementos que los juicios morales, considerados usualmente dentro de la interpretación del observador ideal. Mi posición se apoya en la fuerte relación que Hume establece entre las determinaciones morales básicas propias de la naturaleza humana y la vida social. Hume explica a la sociedad como el producto de ciertas convenciones *tácitas* morales celebradas entre los agentes, cuya finalidad es fortalecer a tales determinaciones básicas y hacer posible la vida gregaria, en sociedad. Estas convenciones son las relativas al respeto a las promesas y la justicia, especialmente el respeto a la propiedad privada. Discuto sus ideas al respecto en el capítulo 5. Ahora bien, aun cuando éstas son reglas morales, los agentes comúnmente no encuentran en ellas mismas motivos para actuar conforme con ellas. Los motivos provienen de ciertos sentimientos muy vivos de utilidad, belleza y benevolencia, pero como éstos no son naturales, entonces los

agentes tienen que proponerse apropiarse de ellos. Es decir, necesitan adoptar una actitud favorable a los motivos morales y, según Hume, esto lo logran aprendiendo de otros agentes y participando activamente en la vida social. Creo que ésta es la plataforma social adecuada para hablar del observador ideal.

Según yo, *el observador ideal* es en realidad un espacio psicológico de deliberación, fuertemente dependiente de diversos criterios sociales. Se forma en cada agente por su relación con otros agentes, especialmente por aquellos reconocidos como autoridad, expertos o guías. En último término, lo que cada agente individual forma mediante un largo proceso de internalización es un conjunto de criterios evaluativos, relativos a la formación de creencias de todo tipo. Estos criterios son básicamente sentimientos informados y dirigidos por la autoridad de otros agentes. Un agente humeano común interioriza tales criterios y la forma en que los *reconoce* en sus deliberaciones y acciones es por medio de una cierta experiencia placentera o dolorosa. No siempre el agente tiene consciencia de tales procesos y, en realidad, lo común es que no tenga ninguna consciencia de ellos. Ocasionalmente, sin embargo, requiere ajustar sus criterios o adquirir nuevos. En estos casos el sentimiento doloroso es un importante elemento motivador. Cuando un agente echa en falta el sentimiento placentero que lo motiva a actuar por deber, puede aun así actuar por deber, quizá hasta fingir que lo hace por placer, pero albergara hacia sí, sostiene Hume, un sentimiento de desaprobación de sí mismo, el cual se expresa como odio hacia sí mismo. Pero esto mismo lo motiva a buscar, mediante la práctica, el sentimiento apropiado de las acciones virtuosas (cf. T 3.2.1.8 SB 479-480; v. sección 6.3). Yo interpreto esta búsqueda como una forma de adquirir los criterios morales correctos, esto es, el punto de vista correcto para evaluar creencias morales y actitudes morales hacia otros agentes y hacia uno mismo. La forma en que el agente *calibra* sus actitudes como genuinamente morales es mediante el reconocimiento de sentimientos placenteros en su deliberación y actuación. Esto muestra, me parece, que la noción de agente moral de Hume es dependiente de convenciones sociales como las mencionadas y muchas otras más particulares, propias de las prácticas morales a las que pertenecen.

Hume analiza, al respecto, la eficacia de *dispositivos* morales de regulación de la conducta como la vergüenza, la alabanza y la fama (T 3.2.12.4 SB 571; v. sección 5.3). Vale la pena enfatizar que la eficacia de estos dispositivos reside en que mediante ellos la conducta de cada agente se somete a la mirada de otros agentes, sus pares. Las prácticas morales humeanas, con base en esto, se pueden caracterizar como una comunidad de agentes en las que cada uno de los miembros es

juez de todos los demás y objeto de juicio de todos los demás. La experiencia moral de cada agente es así sancionada por la aprobación o la reprobación de sus pares. Usualmente, estas sanciones serán recibidas –y otorgadas– de manera expedita. Desde el punto de vista de cada agente, estas sanciones son recibidas, además, mediante sentimientos placenteros o dolorosos, si han sido alabados o reprobados. Desde mi punto de vista, estas ideas de Hume resaltan sobre otras filosofías de la moral porque no da a la *culpa* un lugar preponderante entre los dispositivos morales. Además, tampoco considera que un agente por sí sólo, aislado de la mirada reguladora de sus pares, pueda decidir qué es lo correcto hacer en una situación dada. El agente humeano, en este sentido, no es por sí mismo la fuente de su propia normatividad (también en la sección 5.3 comparo estas ideas de Hume con las correspondientes de la moral kantiana). Esta relación entre los agentes, al interior de una cierta práctica moral, contribuye a la explicación de cómo un agente interioriza las normas y los sentimientos apropiados. Desde mi punto de vista, también contribuye a entender de qué manera el observador ideal adopta los diferentes criterios evaluativos propios de una práctica moral y cómo, con ellos, dirige y regula sus pensamientos y acciones. Por lo demás, Hume aborda todos estos elementos con muy pocos elementos teóricos, todos derivados de su teoría de las ideas, la cual incluye su explicación de cómo las ideas se forman por asociación, esto es, por ciertas determinaciones naturales que dan lugar a las diferentes ideas e impresiones de la mente.

En atención a esta última observación titulé mi trabajo *«La flexibilidad de los sentimientos»: una teoría humeana de la motivación y de los juicios morales* pensando en la fuerza de una de las ideas rectoras de mi investigación: Hume ofrece un argumento convincente sobre la moralidad, los juicios morales y la motivación humana mostrando que si se combinan unos pocos elementos de su teoría de la mente en una explicación o argumento general se puede lograr el mayor alcance explicativo. En efecto, el pensamiento de Hume sobre las impresiones, las ideas, la memoria y la imaginación, así como las determinaciones naturales que las relacionan, logran explicar las creencias sobre cuestiones de hechos, las pasiones, los motivos que tenemos para actuar y los juicios morales. Además, logra explicar cómo todos estos elementos están entre sí relacionados y que, a pesar de su aparente carácter limitado a la naturaleza humana, tiene repercusiones que dan lugar a hacer consideraciones sobre la moral y las convenciones humanas detrás de las reglas morales, como las reglas de justicia, la propiedad privada y las promesas. Creo que la plausibilidad y originalidad del argumento general de Hume es que logra articular de

manera plausible estos elementos propios de nuestra naturaleza con condiciones sociales que los presuponen y los refuerzan. En efecto, para Hume los sentimientos son respuestas valorativas y normativas, “de aprobación”, que se manifiestan de manera flexible y consistente (cf. T 3.3.6.6 SB 617). Las respuestas consistentes se explican porque, para Hume, nuestras impresiones están acompañadas de placer o dolor, y son estas sensaciones los motores básicos de la mente. Los agentes humanos están determinados por naturaleza de manera similar, así que responden de manera consistente, con placer y dolor, a estímulos semejantes, ante diferentes objetos y en diferentes circunstancias, incluyendo, como he dicho, percepciones del mundo sensible o percepciones sociales, relativas a la moralidad y sus reglas. Estas respuestas consistentes son el elemento más básico que explica a la motivación humana.

Mi tesis consta de seis capítulos y sigo en ellos la propia estructura temática del *Tratado*. El primero está dedicado a presentar los elementos básicos de la teoría de la mente y la formación de creencias sobre hechos. En este capítulo examino los elementos básicos que servirán a las teorías de Hume sobre las pasiones, los motivos para actuar y los sentimientos o juicios morales. El segundo capítulo ofrece una interpretación *cognitiva* de la teoría de las pasiones de Hume – sigo aquí las ideas de Anette Baier (1990) y Lili Alanen (2005). Defiendo que las pasiones, a pesar de que Hume a veces afirma algo que parece opuesto a esto, están constituidas por ideas y que tienen algún carácter representacional. En este sentido son cognitivas: en virtud de su carácter representacional estas ideas pueden ser evaluadas como verdaderas o falsas, de acuerdo con la relación con lo que representan (v. sección 2.2.1). Creo que hay mucha evidencia textual en la obra de Hume para defender esta interpretación, la cual, posteriormente, me servirá para explicar que los motivos para actuar presuponen también algún contenido representacional y que, a pesar de que son básicamente sentimientos, son al mismo tiempo estados mentales que ofrecen elementos normativos y valorativos de ciertos objetos. El capítulo tres está dedicado al estudio de la motivación.

Primero, en este capítulo estudio y clarifico la naturaleza de la percepción que hace que el placer o el dolor *nos parezcan* probables o improbables, ciertos o inciertos, etc. Segundo, examino cómo dicha percepción se relaciona con otros estados mentales, como las pasiones o los deseos, para dar lugar a motivos específicos para pensar o actuar. Me concentro en establecer lo más claramente posible la posición de Hume al respecto. Después de esto, caracterizo a estos elementos como una teoría de la motivación en la que la razón desempeña un papel instrumental

al servicio de las pasiones o los deseos, porque creo que es la forma más afín a las ideas de Hume, a pesar de ciertos pasajes donde parece descartar que la razón desempeñe papel alguno en la motivación. Con todo, esta teoría de la motivación ha recibido fuertes críticas de Christine Korsgaard (Korsgaard 1996), las cuales examino. Respondo a sus objeciones desarrollando el modelo sub-humano de la motivación de Bernard Williams (Williams 1993). La propia Korsgaard desarrolla sus críticas considerando las ideas de Hume y Williams, por lo que creo que es viable responder a sus objeciones examinando con cuidado ciertas ideas de Williams sobre el papel de la razón. En todo este tratamiento me concentré en la siguiente pregunta: ¿cómo respondería un agente humeano a una razón práctica motivadora presente a él o ella? Desde luego, una razón práctica puede ser un juicio moral. En el capítulo siguiente, el cuarto, considero la teoría de los juicios morales de Hume.

En este capítulo examino los argumentos por los cuales Hume sostiene que las distinciones morales sobre lo bueno y lo malo, o sobre la virtud y el vicio, son sentimientos morales —y no una forma de creencia o juicio de la razón. En este capítulo presento a las interpretaciones sentimentalistas y del observador ideal de los juicios morales. Respecto de la primera, examino con detalle las dificultades a las que me referí antes y argumento por qué la interpretación del observador ideal responde adecuadamente a ellas. En el capítulo quinto examino el naturalismo moral de Hume.

Me refiero en este capítulo a la concepción de agente de Hume y sus relaciones con la moral. Examino cómo el naturalismo de Hume explica a las prácticas morales sociales con base en las determinaciones naturales estudiadas en los capítulos anteriores, a saber, las relativas a la formación de creencias, las pasiones, los motivos para actuar y los juicios morales. Particularmente, abordo dos problemas centrales. El primero consiste en reconocer que las determinaciones morales naturales son insuficientes para regular la vida social, de modo que se requiere de recursos no naturales, esto es, *artificiales* según los designa Hume, para complementar o reforzar a los primeros. Con esto me refiero a los límites de las virtudes naturales para explicar las regulaciones sociales. Ante esto, Hume argumenta que se requieren de recursos creados socialmente para reforzar a los primeros y explicar las regulaciones sociales. Estos recursos comprenden a las reglas de la justicia, las promesas y el respeto a la propiedad privada. Con todo, Hume observa muy agudamente que tales recursos requieren ya de la vida social, por ejemplo, que los agentes los reconozcan como reglas que limitan su conducta, pero esto sólo es posible si

las propias reglas ya están en funcionamiento. Para responder esto, Hume propone una de sus teorías más interesantes y, me parece a mí, vigentes: la justicia, las promesas y el respeto a la propiedad privada no surgen por acuerdos explícitos, reflexivos, sino por acuerdos tácitos, por actos de convención celebrados sin el auxilio del lenguaje. De este modo, Hume coloca a tales acuerdos en un espacio hipotético previo al surgimiento de las sociedades humanas, al tiempo que los convierte en las condiciones que las hacen posibles. Los agentes siguen tales acuerdos tácitos por la costumbre y el hábito asociados con el placer que obtienen de ello, y no por alguna consideración racional al respecto. De este modo, el escepticismo de Hume sobre la razón también la excluye de participar en las convenciones que dieron origen a las reglas morales básicas. Tales reglas son también expresión del sentimiento moral.

El segundo problema mencionado consiste en cómo los agentes conforman su conducta a las reglas. Su estudio me permite relacionar aspectos relativos a la psicología individual con las reglas de la justicia, las promesas y el respeto a la propiedad privada, pero considerando criterios normativos más generales. Exploro aquí cómo las determinaciones morales naturales se relacionan con las reglas sociales. La filosofía de la moralidad de Hume, por contraste con la kantiana, concibe a los agentes morales preocupados constantemente por el punto de vista de los otros. Esta preocupación tiene un importante papel regulativo, porque los dispositivos morales de control, los criterios normativos más generales mencionados, se refieren a la vergüenza, la alabanza y la fama. Los agentes humeanos no están preocupados por el deber, o lo están sólo en la medida en que a su vez está involucrada la vergüenza, la alabanza o la fama, que serían así los dispositivos morales que regulan en último término su conducta. Específicamente, estos agentes respetan las reglas de la justicia, las promesas y de la propiedad privada, desde un punto de vista moral, pensando y actuando de acuerdo con los dispositivos morales mencionados. En el sexto y último capítulo ofrezco algunas consideraciones finales y hago una valoración del naturalismo de Hume.

Considerando la relación que hay entre los juicios morales y las prácticas morales, en este capítulo me propongo responder la pregunta: ¿qué tan plausible es la explicación de Hume que descansa en los principios generales de la formación de ideas? En este sentido, pretendo valorar su naturalismo moral considerando aspectos relativos a los usos y condiciones de aplicación de los juicios morales (entendidos como valoraciones morales y normas morales). No pretendo detenerme más en el estudio de los posibles objetos a los que se refieren. Es decir, intento hacer

menos meta-ética de la teoría de los juicios morales de Hume y dirigir mi atención en la dirección opuesta –por decirlo así-, la que nos aleja de la teoría sobre la moralidad y nos acerca a las prácticas morales. Esto coincide con otro aspecto observación de Stroud sobre el valor del naturalismo de Hume: “Lo que atrae la atención es la idea misma de una investigación y una explicación empíricas totalmente exhaustivas de por qué los seres humanos son como son y por qué piensan, sienten y se comportan como lo hacen” (Stroud, 325). Creo que para responder esta cuestión es necesario examinar cómo se relacionan elementos como la utilidad y la belleza, característicos de las valoraciones morales de los agentes humanos, con las prácticas morales ordinarias y con los elementos naturales básicos, esto es, las pasiones más básicas (amor, odio, orgullo y humildad). El propio Hume, como digo a lo largo de la tesis, sostiene que el sentimiento moral *requiere* de factores vinculados a las determinaciones naturales básicas y a las condiciones sociales. En este capítulo reúno estos elementos.

Mi conclusión final es que todos estos elementos convergen bien en una interpretación general del pensamiento de Hume que de un lugar central a la interpretación del observador ideal. Según sostengo, el punto de vista social opera en cada uno de nosotros como una especie de plataforma psicológica que sostiene la vida mental de cada individuo agremiado a la sociedad. En otras palabras, nuestra forma de sentir, de pensar, de percibir es una construcción social, no racional, que opera en cada uno de nosotros y que articula todos los elementos mencionados. En buena medida, el punto de vista del observador ideal coincide con el punto de vista social y los criterios sociales.

1. Las determinaciones naturales que explican a las creencias sobre hechos

En este capítulo examino la idea de que *creer*, de acuerdo con Hume, es el producto de las determinaciones naturales de la mente. En particular, me interesa examinar esta idea en relación con las creencias sobre hechos. Hume nos ofrece dos caracterizaciones definitorias de la creencia: “Una creencia se define con más exactitud como una idea vivaz asociada o relacionada con una impresión presente (T 1.3.7.5 SB 96) y: “Una creencia es un sentimiento peculiar que, además de su fuerza y vivacidad, se caracteriza por su *solidez, firmeza o consistencia*” (cf. T 1.3.7.7 SB 629).

Siguiendo a Hume, defiendo que la segunda caracterización es una explicación de la noción de *fuerza* y que los rasgos de solidez, firmeza y consistencia se refieren a la *influencia* que el estado de creer tiene sobre otros estados mentales, como las pasiones y la misma acción. Por esta razón, creo que es un error intentar explicar a los rasgos definitorios de ‘fuerza y viveza’ sin hacer referencia a la solidez, la firmeza y la consistencia del estado de creer, así como a su influencia sobre otros estados mentales y sobre la acción. Además, explicaré cómo estos rasgos son, también, una condición general de la experiencia misma. Las determinaciones básicas de la mente, que dependen fundamentalmente de la naturaleza humana, hacen posible que unos pocos elementos –impresiones e ideas— y unos pocos principios de asociación entre ellos, permiten explicar esta condición. Creo que todo esto se refiere a lo que Hume llamo «la influencia de la creencia sobre la voluntad» y que yo examinaré como *la función práctica* de la creencia (v. sección 1.4 abajo). De acuerdo con esto, las creencias, para Hume, son importantes tanto por su valor cognitivo como práctico. Con ‘valor cognitivo’ me refiero al modo en que las creencias son juzgadas como verdaderas o falsas por las operaciones distintivas de la razón, esto es, por medio de establecer un acuerdo o desacuerdo con relaciones *reales* de ideas, o con la existencia y los hechos *reales* (cf. T 3.1.1.9 SB 458).¹ Con ‘valor práctico’ designo al efecto que las pasiones y los deseos tienen sobre la voluntad para motivar pensamientos (como decisiones) o acciones (los rasgos generales de esta explicación los desarrolla Hume en el libro 2 del *Tratado*, dedicado a las pasiones, v. p.e. T 2.3.). Por último, voy a examinar el argumento de Hume para excluir a la razón

¹ Si bien esta caracterización se encuentra en el “Libro 3”, la elegí para presentar el sentido de cognitivo porque es bastante clara y porque me parece que Hume usa consistentemente este sentido a lo largo del *Tratado*.

como una facultad que pueda explicar a las creencias causales sobre hechos. Este tipo de creencias es quizá el más importante y podría pensarse que, a la luz de la idea de que las creencias tienen valor cognitivo, la razón podría tener un papel crucial en su surgimiento. Sin embargo, Hume argumenta en contra de esto y prueba que aun cuando ciertas creencias puedan surgir de este tipo de investigación deliberada y en apariencia dirigida por la razón, su validez o certeza depende de las mismas determinaciones que explican a las otras creencias sobre hechos. Con base en este argumento, Hume puede sostener la afirmación general de que “Las creencias son producto más del sentimiento que de la parte «cogitativa» de nuestra mente” (T 1.4.1.8 SB 184). Esta afirmación me será muy útil para explicar, en el siguiente capítulo, que las pasiones sean *acerca de algo*, si bien no son enteramente estados cognitivos.

En este capítulo reúno elementos teóricos que emplearé más adelante para hacer una presentación de la teoría de las pasiones de Hume estrechamente relacionada con su teoría de las ideas y de la creencia.

1.1 Las determinaciones básicas de la mente

Hume distingue tres tipos de creencias: 1. Las creencias que son necesariamente verdaderas en virtud de alguna prueba deductiva (las verdades de la aritmética o de la geometría) y dos tipos de creencias sobre hechos o fácticas: 2. Las creencias derivadas directamente de alguna impresión sensible, ya sea proveniente de los sentidos o de la memoria y 3. Las creencias que se refieren a algo que no está directa o inmediatamente presente a los sentidos, aunque sí están relacionadas con alguna impresión sensible, proveniente de los sentidos o de la memoria. Para ejemplificar este último tipo consideremos esta situación: Al llegar a mi calle veo una columna de humo al final de la misma. Aunque no veo directamente el presunto fuego ni el terreno inmediatamente formulo esta creencia: *Se está quemando el terreno baldío al otro lado de mi calle*. Sólo percibo que en esa dirección hay una columna de humo y recuerdo que ahí mismo está un terreno baldío. En este tipo de creencias tengo una impresión sensible relacionada con mi creencia, pero estrictamente no percibo sensiblemente aquello a lo que se refiere la creencia. Esto marca la diferencia con el tipo (2) y, como veremos más abajo, trae varias dificultades explicativas relativas tanto a los procedimientos psicológicos que las explican como al valor epistémico mismo de tales creencias. El principal interés de Hume en el Libro I del *Tratado* parece dirigido a estas dificultades (cf.

Price L. 7, Morris, 77). Tanto su teoría de las ideas como su concepción de la mente parecen dirigidas a explicar este tipo de creencias.

De acuerdo con su modelo de la mente, cualquier suceso mental y cualquier objeto mental se explican por las relaciones entre dos tipos de entidades básicas, las impresiones y las ideas. Las impresiones comprenden «todas nuestras sensaciones, pasiones y emociones», mientras que a las ideas las concibe como «imágenes débiles de las impresiones en nuestro pensamiento y razonamiento». La diferencia entre impresiones e ideas se reduce al distinto grado «de fuerza y vivacidad con que inciden sobre la mente» (cf. T 1.1.1.1 SB 1). Una idea está en la mente con un grado menor de fuerza y vivacidad que una impresión. Hume piensa que esta distinción es fácil de apreciar: es la diferencia entre percibir sensorialmente y recordar esas percepciones, por ejemplo, es la distinción que hay entre saborear de hecho algo y recordar ese sabor. Saborear una manzana es una impresión sensible y el recuerdo del sabor es una idea. Podemos aceptar esta distinción sin objeciones (Stroud, p.e., señala que «es claro qué quiere decir Hume con ella y por qué intenta trazarla así», cf. Stroud, 35). El empirismo defendido por Hume descansa en esta idea, la cual le permite rechazar al innatismo –la idea de que algunas de nuestras ideas no se originan en la experiencia sensible- y sostener que todas nuestras ideas se originan mediante alguna percepción sensible o algún sentimiento (cf. T 1.1.1.12 SB 7; Stroud, 54; Broughton, 51-52). Si todos los contenidos de la mente se originan en impresiones sensibles, entonces podemos suponer que las ideas son derivadas de ellas. Si queremos darle a un niño, dice, la idea de rojo o naranja, de lo dulce o lo amargo, le presentamos objetos apropiados, le mostramos aquellos colores o le damos a probar aquellos sabores. Por ejemplo, no parece que nadie pueda adquirir la idea del sabor del rambután (un fruto originario de Malasia) sin probarlo directamente. Hume sostiene, en efecto, que «las ideas dependen de las impresiones» (cf. T 1.1.1.8 SB 4-5). Éste es «el principio de prioridad de las impresiones sobre las ideas» y parece fundamentalmente correcto. Sin embargo, también parece verdad que podemos formarnos alguna idea del sabor del rambután mediante comparaciones, por ejemplo, si alguien que lo ha probado trata de describirlo para nosotros mencionando sabores o texturas que se consideran semejantes. Aun así la idea que lleguemos a formarnos será dependiente de ciertas impresiones sensibles *propias*, como el recuerdo del sabor que nos describen alojado en nuestra memoria y con el cual comparamos la descripción que nos ofrecen. Con esta observación podemos apreciar las relaciones básicas que hay entre la memoria y la imaginación.

La descripción del sabor que nos presenten activara nuestra *imaginación* y llegaremos a concebir de alguna manera el sabor o la textura que nos describen, porque la idea del sabor es, de algún modo, una *copia* de la impresión sensible del sabor (cf. T 1.1.1.6 SB 4). Este proceso que relaciona, en la memoria, a las impresiones sensibles con sus correspondientes ideas es la primera determinación básica de la mente que presenta el *Tratado* y se suele identificar como el «principio de la copia». Según éste, la primera manifestación en la mente de cualquier idea *simple* se deriva de alguna impresión *simple*, con la cual se corresponde y a la cual representa exactamente (cf. T 1.1.1.4-6 SB 3-4). Hume parece aceptar que la distinción entre ideas simples y compuestas es un supuesto necesario para explicar otras determinaciones de la mente de las que tenemos conocimiento directo.

Sin embargo, desde el punto de vista de Hume, el problema con las ideas abstractas y aun con las ideas simples consiste, brevemente expuesto, en reconocer que como nunca percibimos directamente al azul o al frío entonces no tenemos claro cuál pueda ser el origen de estos tipos de ideas. Nuestras percepciones al respecto son siempre *complejas* en el sentido de que percibimos una manzana roja o el manar de la roja sangre o *sentimos* frío al levantarnos de la cama o al morder una paleta de hielo. En estas percepciones reconocemos en conjunto aspectos como la forma de los objetos, su textura, sus matices, su localización (sentir frío en todo el cuerpo es distinto a la sensación local de frío proveniente de morder una paleta), etc. Hume se limita a explicar el origen de estas ideas recurriendo al principio de la copia: esas ideas son copias fieles de sensaciones simples. Pero considerando cómo he planteado el problema parece que esta respuesta es por lo menos polémica, porque, en efecto, parece obvio que nunca tenemos percepciones sensibles simples en el sentido en que parece requerir la explicación de Hume. Dos alternativas parecen plausibles y yo creo que son complementarias entre sí: por una parte, tendríamos que aceptar que la simplicidad de las sensaciones e ideas no es analizable por ningún método que conozcamos para estudiar a las ideas abstractas o a nuestras costumbres lingüísticas (cf. Baier 1991, 35). Así, las ideas simples serían primitivas y necesarias para explicar a las operaciones de la imaginación. La otra alternativa es *interpretar* que Hume realmente no sostiene por completo una teoría *representacionalista* de las ideas, o al menos que no acepta todos los elementos de este tipo de teorías. Una teoría de este tipo sostendría sin reserva alguna que las ideas son copias de las impresiones y que, a su vez, éstas son representaciones fieles de objetos del mundo natural. De acuerdo con Fate Norton, Hume conocía bien las implicaciones escépticas de este tipo de teorías

y, dado su desinterés en las conexiones entre las ideas y las impresiones de sensación,² así como su idea de que todos los objetos de la mente son siempre percepciones, dice que debemos concluir que “él no pensó que éstas fueran, en un sentido cardinal, representativas de objetos –ni las impresiones ni las ideas semejan objetos” (cf. Fate Norton 2003, 7). La cautela de Hume al decir que las impresiones de sensación surgen de causas desconocidas y que tales causas son por completo inexplicables para la razón humana apoya, me parece, esta interpretación de Fate Norton (cf. T 1.1.2.1 SB 7-8 y T 1.3.5.2 SB 84). Hume no niega que pueda haber alguna relación de representación entre los objetos del mundo natural y las impresiones de sensación supuestamente correlativas de los mismos, pero tampoco afirma que la haya. De este modo, Hume no parece comprometido con la afirmación de que las impresiones de sensación *representan* a los objetos del mundo natural, pero sí con que las ideas son representaciones fieles de las impresiones de sensación. Este matiz respecto de la noción de representación que puede derivarse de su teoría de las ideas será relevante para explicar, más adelante, de qué manera las pasiones son intencionales o cómo es que son *acerca* de ciertos objetos (v. cap. 2, p. 49). Lo que aprendemos de estas nociones sobre las ideas simples y complejas nos enseña mucho de las determinaciones básicas de la mente donde intervienen la memoria y la imaginación.

Hume describe a la operación de la memoria como ocupada en preservar copias de las impresiones sensibles y conservar en algún grado su vivacidad original, mientras que la imaginación combina a las ideas con relativa libertad (cf. T 1.1.3.1 SB 8). Estas operaciones *presuponen* necesariamente a la distinción entre ideas simples y complejas: Cuando pienso en un fruto como el rambután y en su color, en mi mente se avivan ciertos recuerdos relativos a estas ideas y la imaginación juega con cierta libertad con ellas, por ejemplo, imaginando que es de color amarillo. Así, separa de las impresiones sensibles la idea de amarillo y la asocia en la idea compleja el rambután es un fruto amarillo. Esta correlación parece necesaria para *imaginar*, por ejemplo, que el rambután *necesariamente* tiene algún sabor o color. Cualquier cosa que imaginemos al respecto tiene que combinar alguna idea proveniente de la memoria y una operación de la imaginación por medio de la cual surge una idea compleja. Las determinaciones naturales de asociación más básicas son posibles precisamente porque operan sobre las ideas simples: “Como todas las ideas simples pueden ser separadas por la imaginación y unidas de

² Como vimos, Hume se limita a señalar que «no cree que sea necesario gastar muchas palabras para explicar la distinción entre impresiones e ideas», cf. T 1.1.1.1 SB 1.

nuevo en la forma que a ésta le plazca, nada sería más inexplicable que las operaciones de esta facultad si no estuviera guiada por algunos principios universales que la hacen, en cierto modo, conforme consigo misma en todo tiempo y lugar” (T 1.1.4.1 SB 10). La imaginación combina diversas ideas y las integra en complejos que, como cualquiera puede observar, son bastante libres pero no por completo arbitrarios, *i.e.*, no podemos imaginar nada que «implique una absoluta contradicción o absurdidad» (cf. E §2.2 SB 18; T 1.3.3.3 SB 80). Esta limitación apunta a que la imaginación actúa regulada por ciertos principios o reglas. Por ejemplo, a pesar de la libertad de la imaginación, no solemos pensar que el sabor de una fruta mida tres metros de largo o que su color sea el pájaro que veo volando por la ventana. La imaginación relaciona ideas de acuerdo con regularidades o principios de atracción que «la hacen conforme consigo misma en todo tiempo y lugar». Hume no nos ofrece una prueba última sobre las ideas simples, como hemos visto, ni sobre los principios de asociación, pero parece claro que su prueba general respecto de cómo se relacionan ambos grupos de elementos descansa en la observación de que, *de hecho*, la experiencia parece constituida de correlaciones de ambos tipos de elementos. Según Hume, las determinaciones básicas que constituyen a la atracción entre ideas son la contigüidad en tiempo y lugar, la semejanza y la causalidad (cf. T 1.1.4) y, como decía arriba, su origen es incierto y se remonta a la propia naturaleza humana. Hume no ofrece ninguna prueba argumental a favor de estas relaciones. Se limita a observar ciertos hechos, como que en la constante revolución de nuestras ideas la imaginación pasa fácilmente de una idea a cualquier otra que le sea parecida y que esto es suficiente para constituir un vínculo de asociación entre ambas ideas. La contigüidad de dos objetos permite a la imaginación relacionarlos, como ver una columna de humo al otro lado de mi calle y relacionarlos en la idea de que *ahí* hay un incendio. Finalmente, la relación de causa y efecto produce una impresión muy fuerte en la mente, *parecida* a una convicción (cf. T 1.1.4.2-5 SB 11-12). Por ejemplo, mi creencia de que hay un incendio al otro lado de la calle presupone, sin más, que el fuego en el terreno *causó* a la columna de humo. En algunas circunstancias, con esta creencia puedo tomar medidas prácticas importantes las cuales, al menos, acusan el grado de convicción que le concedo. Estas tres relaciones de asociación entre ideas son un rasgo del naturalismo humeano que explica la determinación cuasi-mecánica detrás del surgimiento de los distintos estados mentales, incluyendo el de tener una convicción. Hume considera que su descubrimiento de estas relaciones entre ideas y el principio de asociación que las reúne es su contribución más importante a la filosofía (cf. T *An abstract* 35 SB 661-662).

Respecto de la ciencia de la mente considera que su importancia es análoga al descubrimiento de la propia atracción para la ciencia de la naturaleza (cf. T 1.1.4.6 SB 12-13).³ El principio de asociación entre ideas, además, nos deja ver que la memoria y la imaginación intervienen prácticamente en todas las operaciones mentales.

En efecto, aun cuando no siempre sean mencionadas, un supuesto que se deriva de estas observaciones sobre las operaciones de la mente es que la memoria y la imaginación intervienen en cualquier proceso mental, tanto en el ámbito teórico como en el práctico (volveré numerosas veces a esta tesis en los siguientes capítulos). En el modelo de la mente de Hume las determinaciones básicas provenientes de dichas facultades están prácticamente detrás de todos los estados mentales, incluyendo los sentimientos, las pasiones, el estado creer y aun el de juzgar (cf. Kemp Smith 477; Stroud, 25; Merrill, 267).⁴ La función principal de tales determinaciones parece ser la de transmitir la vivacidad original de las impresiones sensibles originales a cualquier otra idea relacionada o asociada con ella (cf. Kemp Smith, 377). Hume recurre constantemente a esta idea para explicar la fuerza con que ciertos estados mentales, como las pasiones o las propias creencias, inciden sobre la mente. La forma en que la imaginación logra hacer esto es variando el *modo* en que una idea es concebida por la mente (cf. T 1.3.7.2 SB 95).⁵ Con estos antecedentes examinemos cómo caracteriza Hume a la creencia.

³ Hume con su principio de atracción de las ideas pretende emular al propio principio de atracción universal de Newton, tal como explica Stroud: “Tal como la ley de la gravitación universal de Newton da razón del movimiento y las posiciones subsecuentes de todas las partículas físicas del universo, la concepción formal de Hume supone que lo que llama “el principio de la asociación de ideas” ha de dar razón de todos los fenómenos mentales y psicológicos explicando cómo y por qué las diversas percepciones llegan a estar «presentes a la mente»” (Stroud, 21). Cf. también E §1 SB 14-15, donde Hume también espera que semejantes a los descubrimientos de Newton sobre la fuerza y leyes que gobiernan a los planetas se descubran “los poderes mentales y su estructura”.

⁴ Ésta es una de las ideas centrales de la pionera interpretación de Kemp Smith sobre Hume. Según Kemp Smith la imaginación opera sobre las ideas de manera análoga a como la simpatía opera sobre las pasiones o sentimientos (v. también 222, aunque es una idea que elabora a lo largo del libro). Merrill también observa que para Hume la simpatía es la fuerza básica que opera comunicando y transformando ideas en sentimientos y en impresiones. Me parece que estas ideas equivalen, en general, a caracterizar como flexible al modelo de la mente de Hume.

⁵ Desde luego, Hume no presenta la distinción entre estados mentales en términos de «actitudes proposicionales», aunque su tratamiento del tema hace pensar que podría tratarse del mismo concepto (cf. Kemp Smith, 222; Merrill, 59-60; Stroud 114 y 236; Bennett 2001, 224) A pesar de que sea plausible esta interpretación, creo que las ideas de Hume son más interesantes, por tener una idea más amplia de las características de los estados mentales. Defenderé más adelante, en la presentación de su teoría de las pasiones, que para Hume no todos los estados mentales se pueden describir como actitudes proposicionales. En particular, siguiendo a Baier y Alanen, sostendré que las pasiones no tienen contenido proposicional, aunque sí son intencionales (v. infra p. 49 y ss). Tal vez pueda ahora decir algo breve al respecto. Imaginar, por ejemplo, que el rambután sabe a mango no es imaginar realmente el sabor distintivo del rambután, sino es meramente dejarse llevar por la capacidad que tenemos para pensar por medio de analogías. Podría decir, también, que sabe a pollo o frío aluminio, etc. Si a propósito de esto decimos, por ejemplo, ‘*creo que el rambután sabe a mango*’ parece que formulamos una expresión con contenido proposicional, pero por lo que digo

1.2 Las impresiones sensibles y las creencias sobre hechos derivadas directamente de ellas

Según Hume, la percepción de una impresión sensible da lugar de manera directa e inmediata a una creencia. Lo que marca al estado de creer, en este caso, es la fuerza misma de la percepción sensible:

...la *creencia* o *asentimiento*, que acompaña siempre a la memoria y los sentidos, no es otra cosa que la vivacidad de las percepciones que presentan esas facultades... En este sentido, creer es sentir una inmediata impresión de los sentidos, o una repetición de esa impresión en la memoria (T 1.3.5.7 SB 86).

Estas creencias se originan *directamente* en la fuerza o vivacidad de la impresión y por ello no representan, para la teoría de Hume, ningún problema de explicación.⁶ Fate Norton considera que este tipo de impresiones ofrece el modelo de la forma de creer que Hume pretende, en general, explicar (cf. Fate Norton 2000, I27-28). De acuerdo con este modelo, el componente distintivo de la creencia es una impresión presente y el estado de creer parece identificarse plenamente con la percepción fenoménica de la fuerza de la impresión. Según Fate Norton, el proceso que explica a las creencias sobre hechos del tipo (1) es mecánico y causal.⁷ En efecto, no parece que se requiera de mucho más para explicar a este tipo de creencias: la percepción sensible del humo me hace creer que hay humo. Pocas veces, y solo cuando tenemos razones para dudar de nuestras percepciones, dudamos de entrada de lo que percibimos. Ahora bien, como por regla general las creencias de este tipo se *refieren* a impresiones sensibles entonces tendremos que aceptar que si cabe hablar de sus cualidades intencionales o de representación, éstas cualidades tendrían que entenderse dentro de los límites mecánicos y causales de esta concepción de la referencia. Como

parece que esta expresión no es proposicional porque no puedo estar hablando del sabor del rambután, sino de una cierta analogía. De modo que el contenido proposicional no parece el de una creencia, sino, en todo caso, el de una mera concepción. Pero posiblemente solo se está sucumbiendo al *embrujo* del lenguaje y su capacidad de formar analogías sin tener en la mente un pensamiento proposicional.

⁶ Dada la teoría de Hume sobre la creencia, según la cual una creencia es una idea fuerte y vivaz asociada con una impresión, no parece que tendría que haber dudas sobre su posición respecto de las creencias empíricas de este tipo, sin embargo, autores como Price niegan explícitamente que para Hume éstas sean precisamente creencias: "...statements describing a present sense-impression or a present impression of reflection, in Hume's terminology; for example 'This looks red to me now' or 'I am now feeling frightened'. Here again there is no room for belief, on Hume's view. We are just directly aware of these impressions at the moment when we experience them" (L7 §3). Price, en parte, fundamenta en este pasaje su crítica de que la teoría de la creencia de Hume es demasiado estrecha. Sin embargo, me parece que no tiene razón porque las percepciones sensibles son la evidencia de las creencias respectivas. Estas percepciones y los enunciados descriptivos asociados, cuando las hay, constituyen a la creencia y son así muy distintos de una mera concepción.

⁷ Hume mismo, en efecto, considera que la mente tiene una estructura causal: "La idea verdadera de la mente humana es considerarla como un sistema de diferentes percepciones o existencias, a las que estamos enlazados por la relación de causa y efecto, y que mutuamente se producen, destruyen y modificación unas a otras" (T 1.4.6.19 SB 261).

veremos en la siguiente sección, esto deja muy poco espacio para la intervención de la razón en la formación de creencias y de juicios (cf. Alanen 2006, 179). Hume mismo considera que estos procesos causales, que identifica como un tipo de atracción entre ideas, son misteriosos y «descansan fundamentalmente en las cualidades originales de la naturaleza humana» (cf. T 1.1.4.6 SB 13). Creo que esta intuición de Hume sobre la referencia es básicamente correcta. La naturaleza humana está configurada de maneras determinadas para percibir al mundo de una cierta manera y, hasta donde podemos saber, de manera muy distinta a como lo percibe un gato o un murciélago. Consideremos ahora al tipo (3) de creencias sobre hechos. Este tipo de creencias se caracteriza precisamente por no tener esa relación directa con alguna impresión sensible, sin embargo, según Hume aun estas creencias se explican por mantener alguna relación indirecta con alguna impresión sensible.

1.3 Las creencias sobre hechos y la «definición» de creencia

De acuerdo con Hume, las impresiones y las ideas están presentes en todas las operaciones de la mente y son constitutivas, de una forma o de otra, de todos los estados mentales. Las ideas y las impresiones, según Hume, solo se distinguen por su distinto grado de viveza, de modo que concebir o pensar una idea es, en un sentido importante, el mismo tipo de pensamiento que tener una impresión. Ahora, la viveza de las percepciones de sensación causa, como hemos visto, directamente un cierto estado mental que Hume identifica con una creencia sobre aquello que percibimos sensiblemente. La viveza asociada con la impresión parece conducir a la mente a *conceder* su asentimiento a lo que percibe directamente. Sin embargo, en el tipo (3) de creencias sobre hechos no está presente esta relación directa con una impresión, de modo que las ideas relativas a hechos, formadas por la imaginación a partir de impresiones sensibles directas y provenientes de la memoria, son meros pensamientos que carecen de la viveza de creencias del tipo (2). A pesar de esto, *tenemos* creencias sobre hechos del tipo (3). Es decir, los elementos presentes en los estados mentales que son las creencias del tipo (3) son *exactamente* los mismos que encontramos en un estado mental producido por la imaginación, en una suposición o en cualquier pensamiento relativo a hechos empíricos que no sea una percepción directa. ¿Qué es, entonces, lo característico y definitorio de las creencias del tipo (3)? Hume aborda esta cuestión distinguiendo entre meramente tener un pensamiento o una concepción acerca de algo y tener una creencia acerca de algo.

La mera concepción de una idea no agrega nada al pensamiento de la misma. Además, Hume añade que de la mera consideración o pensamiento de una idea no se deriva en absoluto la idea de que tal cosa exista. Hume lo reconoce así: “Es también evidente que la idea de existencia no se diferencia en nada de la idea de un objeto y que, cuando concebimos una cosa como existente después de haberla concebido sin más, no hacemos en realidad adición o alteración alguna de nuestra idea primera” (T 1.3.7.2 SB 94).⁸ Imaginar, en efecto, que existe tal o cual objeto, como una montaña de oro o a Pegaso cubierto de flores no es más que una idea de esos objetos. Hume parece indicar con esto un aspecto *normativo* asociado con la idea de creer, porque sugiere que no estamos autorizados, por alguna razón que no establece, a sostener la idea de existencia de algo con base en la mera concepción de ese algo. Su punto parece ser que sólo una creencia sobre hechos permite atribuir existencia a una cierta idea, en virtud precisamente de la relación adecuada con alguna impresión presente. Tener una creencia sobre un hecho es, entonces, la expresión de un *modo* de pensar a la idea, mediante el cual se afirma la existencia de un objeto que no está presente de manera inmediata a nuestra percepción, pero que tiene alguna relación causal adecuada con alguna impresión sensible o de la memoria. El modo de pensar a las ideas es lo que distingue a la mera concepción de la creencia (cf. T 1.3.7.2 SB 94-95). En consecuencia, lo que se requiere para explicar a la creencia es que además de las variaciones provenientes de la imaginación, tenga lugar «un cambio en el *modo* de concebir a las ideas» (cf. T 1.3.7.4 SB 96). Desde el punto de vista del agente, el modo peculiar de la creencia se aprecia por un cierto rango de intensidad fenoménica. De acuerdo con Hume, la única forma en que podemos explicar el modo distintivo de concebir a una idea para que devenga una creencia es incrementando su fuerza. Explica esta variación mediante una analogía con los colores: “Si deseáis variar de algún modo la idea de un objeto particular, lo más que podéis hacer es incrementar o disminuir su fuerza o vivacidad. [...] Lo mismo sucede con los colores. Un matiz determinado de color puede adquirir un nuevo grado de viveza o brillo, sin ninguna otra alteración.” (cf. T 1.3.7.5 SB 96). Por estas razones, las creencias en general, para Hume, o más precisamente el estado de creer, se caracteriza por su fuerza peculiar: (1) “Las creencias son *ideas* que han ganado en fuerza y vivacidad (T 1.3.6.7 SB 86). En el caso de las creencias de tipo (3), *i.e.* aquellas donde no hay una impresión sensible involucrada directamente, la creencia surge de una idea que se ha asociado con alguna

⁸ Bennett señala, con base en esta cita, que Hume no sostiene ninguna idea de existencia (cf. Bennett, 226). Hume rechaza así a los argumentos ontológicos que, en efecto, pretenden derivar la existencia de Dios del análisis de las propiedades del concepto ‘Dios’ (cf. T 1.3.7.2 SB 94, siguientes líneas de la cita anterior).

impresión sensible, mediante un proceso que tiene lugar de manera espontánea y se debe *exclusivamente* a las determinaciones naturales de la mente. Por medio de estos procesos la viveza o fuerza de una impresión sensible se *transmite* a la idea que será, bajo este modo de ser percibida, una creencia. Hume identifica a este proceso como uno de los principios más importantes para producir cambios en el modo en que una idea es percibida y los correspondientes cambios de estado de la mente:

Desearía establecer como máxima general de la ciencia de la naturaleza humana que siempre que una impresión cualquiera llega a sernos presente no sólo lleva a la mente las ideas con las que está relacionada, sino que comunica también a estas últimas parte de su fuerza y vivacidad. Todas las operaciones de la mente dependen en gran medida, cuando ésta las realiza, del modo en que estén dispuestas... (cf. T 1.3.8.2 SB 98).

Con todo, este principio parece actuar en la formación de creencias en conjunción con las relaciones naturales operadas por la imaginación. La semejanza, por ejemplo, opera así: La percepción de un objeto sensible, sea que lo percibamos directamente por medio de una impresión o por medio de una idea de la memoria, comunica su fuerza, por semejanza, a la idea de otro objeto semejante relacionado con él. Éste es precisamente el proceso que hace surgir a las creencias sobre hechos del tipo (3):

Los objetos sensibles tienen siempre más influencia en la fantasía que cualquier otra cosa, y llevan prontamente esta influencia a las ideas con ellos relacionadas y que se les asemejan. De estas prácticas, y de este razonamiento, inferiré tan sólo que el efecto de la semejanza: avivar la idea, es muy común, y que como en todos los casos deben concurrir una semejanza y una impresión presente... (T 1.3.8.4 SB 100)

En consecuencia, las creencias sobre hechos del tipo (3) se pueden explicar por la concurrencia de estos dos principios, el de transmisión que señala simplemente que una impresión sensible comunica su fuerza a cualquier idea con la que se relacione y el principio de semejanza, que opera a través de la imaginación relacionando, precisamente, con una impresión sensible de un objeto una idea semejante. Sobre esta operación de la imaginación actúa el principio de transmisión. El mismo efecto se observa en la contigüidad.

La operación de la contigüidad entre ideas de objetos se realiza de manera inmediata por la mente, pero este proceso no da lugar a creencias. Cuando pienso en un paseo que he hecho con amigos surgen en mi mente diversas ideas relativas a ellos. Recuerdo, por ejemplo, donde comimos y este pensamiento me hace pensar en un platillo que quiero preparar. En mi mente estos pensamientos son contiguos pero no dan lugar a creencias. Sin embargo, si en ese tren de ideas se hace presente la percepción de un objeto real, el proceso gana inmediatamente en vivacidad.

Podría pensar en el platillo que quiero preparar y esta percepción me pone en contacto con el recuerdo de los ingredientes en la alacena. De esta impresión sensible, muy «fuerte» por su inmediatez, genero la creencia de que tengo todo para prepararlo, excepto ajo. De este modo, la fuerza de la percepción sensible se comunica a las demás ideas y llego a una idea determinada, que es más fuerte que la mera sucesión de ideas en mi mente (cf. T 1.3.8.5 SB 100). Estos ejemplos son suficientes para establecer que es la relación con una impresión sensible lo que modifica el modo en que una idea es percibida por la mente. Si este proceso conduce, por transmisión, a un incremento en la fuerza de la idea, es decir, si la intensidad fenoménica con la que la mente la percibe se incrementa, entonces surge una creencia. Estas determinaciones de la mente explican, ahora, la afirmación definitoria de las creencias: “...una opinión o creencia puede definirse con mayor exactitud como IDEA VIVAZ ASOCIADA O RELACIONADA CON UNA IMPRESIÓN PRESENTE” (T 1.3.7.5 SB 96, énfasis de Hume). Sin embargo, esta afirmación requiere matizarse, porque por sí misma conduce a algunas dificultades. Como veremos enseguida, Hume mismo ofrece una precisión importante para explicar el sentido especial de ‘fuerza’. Las dificultades de la definición sin matizar, citada arriba, se pueden apreciar en ciertas observaciones hechas por Bennett.

Bennett crítica a la teoría de Hume por identificar a las creencias sobre cuestiones de hecho con este peculiar sentimiento fuerte y vivaz, lo que, según él, da lugar al problema siguiente: “si la vivacidad de las ideas visuales es su brillantez, entonces mi estado mental cuando creo que la pared detrás de nosotros es de un azul cielo pálido tiene que parecerse al tuyo cuando meramente contemplas la posibilidad [sin creerla] de que sea azul cielo brillante” (Bennett, 229). La objeción de Bennett sería justa si *sólo* la fuerza y vivacidad distinguiera a las creencias de hecho de una mera concepción (*i.e.*, la diferencia entre meramente concebir una idea y creerla), pero en realidad es la relación con una impresión sensible lo que marca la diferencia entre creer y meramente concebir. Bennett parece, entonces, cometer el error de identificar a una idea vívida con una imagen vívida (cf. Price, L7),⁹ pero, por lo dicho antes, podemos sostener ahora que la viveza de una idea –incluso si entendemos literalmente este atributo– no es suficiente para dar lugar a una creencia sobre cuestiones de hecho. Por alguna razón una idea (en términos de Hume) podría ser

⁹ Ciertas presentaciones de la conclusión general del argumento general de Hume sobre la creencia parecen estar detrás de esta interpretación: “De este modo, se ve que la *creencia* o *asentimiento* que acompaña siempre a la memoria y a los sentidos, no es otra cosa que la vivacidad de las percepciones que presentan esas facultades...” T 1.3.6.7 SB 86. El énfasis que Hume pone por medio de “no es otra cosa que la vivacidad” podría hacernos olvidar la dependencia que tal vivacidad tiene con las impresiones presentes, ya sean de la memoria o de los sentidos.

fuerte y vívida –por ejemplo, porque el agente padece de fiebre alta o consumió alguna droga, etc. (cf. T 1.1.1.1 SB 2)– e incluso afectar a la conducta, sin que estrictamente este presente una creencia en tal estado vivaz, porque el requisito imprescindible para definir a la creencia no es la fuerza si no la relación con alguna impresión sensible. Pero Hume, como anticipé un párrafo antes, matiza su definición caracterizando a la noción de fuerza en un sentido especial (cf. Morris, 82).

Hume mantiene que las creencias son un sentimiento peculiar, caracterizado por su fuerza, pero agrega otros rasgos: “Una creencia es un sentimiento peculiar que, además de su fuerza y vivacidad, se caracteriza por su *solidez, firmeza o consistencia*” (cf. T 1.3.7.7 SB 629, énfasis de Hume). Hume califica así a una idea más fuerte –una creencia– por su mayor solidez, duración o persistencia o influencia que una mera idea (cf. Kemp Smith, 223; Price 7; Morris, 82). Es verdad que Hume mismo no parece satisfecho con su caracterización definitiva del sentimiento de creer,¹⁰ aunque para él la importancia del mismo, o al menos de su explicación sobre ellas, apunta a sus relaciones con otros estados mentales, como las pasiones. Hume desarrolla estas ideas bajo el tema de *la influencia de la creencia* y con él parece referirse a cómo la creencia influye sobre las pasiones y la experiencia misma (cf. Broughton, 45; Schwitzgebel 2006).¹¹ En mi investigación designaré, para abreviar, a tal influencia como «la función práctica de la creencia». En efecto, creo que la mejor manera de entender a la caracterización del estado de creer en términos de solidez, duración e influencia es mediante la relación que guarda con otros estados mentales y, particularmente, por la manera en que es constitutivo de la experiencia misma. Esta particularidad me parece importante porque muchos otros estados mentales se relacionan con las pasiones. En atención a esto, sostengo que el estado de creer, para Hume, constituye a una cierta experiencia y con ello, su capacidad para influir sobre las pasiones y la acción nos pone en condiciones de ofrecer explicaciones relevantes, en general, para la teoría de la acción y de la moralidad (examinaré estas explicaciones en los capítulos siguientes). Con todo, reconozco que

¹⁰ Hume reconoce que no ha capturado correctamente el rasgo distintivo de las creencias, si bien la propia teoría le parece correcta: “...tengo que confesar que en este caso [explicar la operación mental de que surge la creencia] me resulta considerablemente difícil y que, aun cuando creo entender perfectamente el asunto, me faltan términos para expresar lo que quiero decir. ...reconozco que es imposible explicar a la perfección este sentimiento o modo de concebir [la creencia]. Es verdad que podemos usar palabras que lo expresen aproximadamente, pero su nombre propio y verdadero es el de *creencia*, término que todo el mundo entiende suficientemente en la vida común”, cf. T. 1.3.7.7 SB 628-629).

¹¹ Broughton, con todo, se refiere a la ‘viveza’ de una idea: según él «la única forma de entender a la viveza de una idea es por su capacidad para influir sobre las pasiones». Sin embargo, con su interpretación reconoce esta relación de la creencia con la pasión.

si es plausible defender estas relaciones e influencia del estado de creer entonces el sentido de ‘práctico’ que podemos elaborar con base en las ideas de Hume es más amplio que el mero sentido relativo a la deliberación y la toma de decisiones. Consideraré en seguida las condiciones que podrían hacer plausible aceptar este sentido amplio.

1.4. La influencia o función *práctica* de la creencia

El conjunto de las determinaciones naturales que hemos considerado permiten afirmar que no podemos evitar creer. Tampoco podemos creer en algo por voluntad propia (cf. Stroud, 108). El acto de creer no es así un acto de la voluntad, pero sí es un rasgo definitorio y vital de nuestra naturaleza: “creer, o juzgar, es una determinación absoluta e incontrolable de nuestra naturaleza, tan natural como respirar o sentir” (T 1.4.1.7 SB 183). El estado de creer forma parte del conjunto de determinaciones naturales cognitivas, si bien, como hemos visto, para Hume es también un tipo de sentimiento. Por su parte, los sentimientos constituyen otro grupo de determinaciones básicas de nuestra naturaleza. A la par que las determinaciones cognitivas que operan sobre las impresiones y las ideas, Hume introduce, desde las primeras líneas del *Tratado* donde presenta su teoría de las ideas, a ciertos sentimientos asociados con las impresiones sensibles. Estos sentimientos marcan fenoménicamente a las impresiones sensibles de manera distintiva:

Las impresiones pueden ser de dos clases: de SENSACIÓN y de REFLEXIÓN. La primera clase surge originariamente en el alma a partir de causas desconocidas. La segunda se deriva en gran medida de nuestras ideas, y esto en el orden siguiente: una impresión se manifiesta en primer lugar en los sentidos, y hace que percibamos calor o frío, placer o dolor de uno y otro tipo. [...] Esta idea de placer o dolor, cuando incide a su vez en el alma, produce las nuevas impresiones de deseo y aversión, esperanza y temor, que pueden llamarse propiamente impresiones de reflexión, puesto que de ella se derivan (T 1.1.2.1 SB 7-8, énfasis de Hume).

El pasaje parece implicar que toda impresión de sensación viene acompañada de un sentimiento placentero o doloroso. En la amplia discusión sobre la fenomenología de las pasiones en el *Libro II* del *Tratado*, que requiere un estudio pormenorizado del placer y el dolor, Hume hace explícita esta implicación: “...no aparece ante los sentidos objeto alguno, ni se forma imagen en la fantasía, que no estén acompañados de alguna emoción...” (cf. T 2.2.8.4 SB 373). Cuando experimentamos frío o calor, un pellizco o un machucón, un color o una textura, percibimos simultáneamente placer o dolor, agrado o desagrado. Podría parecer cuestionable que percibir un color o una forma sea placentero o doloroso, pero esta cuestión surge, me parece, porque no tenemos memoria directa de la primera vez que percibimos algo azul o vimos un círculo o tocamos algo rugoso. Cabe suponer, por mor del argumento, que la primera experiencia de una impresión sensible dada está

acompañada de cierto placer o dolor que marcan a dicha experiencia y que esto propicia ciertas respuestas emocionales al respecto –según veremos, esto es de importancia para la teoría de las pasiones de Hume (v. infra §2.2). En el pasaje citado arriba, Hume explica que en casos como estos «la costumbre» ha borrado la marca fenomenológica del placer o dolor, haciendo que la sensación misma se confunda con el objeto o idea (“...aunque la costumbre nos pueda llevar a pasar por alto esta sensación y a confundirla con el objeto o idea...”). Hume plantea su teoría de las ideas suponiendo en efecto que cada una de nuestras ideas se origina en alguna impresión sensible primordial, de la que guardamos una copia. Esta impresión primordial parece siempre asociada con el placer o el dolor. Mi intuición al respecto es que esta especie de memoria sentimental es relevante para explicar, de alguna manera, porqué preferimos un color a otro, ciertas formas a otras, etc. Por otra parte, Hume no ofrece, me parece, una prueba argumentativa directa para defender esta relación del placer o dolor con las impresiones sensibles. Al respecto parece seguir una estrategia común en el *Tratado*: sostiene ciertas afirmaciones porque parecen la mejor explicación para ciertos hechos. En este caso, el placer o dolor primordial parecen necesarios para explicar adecuadamente el hecho de que naturalmente buscamos el placer y rechazamos el dolor: “La mente tiende por un instinto *original* a unirse al bien y evitar el mal, aun en el caso de que éstos sean concebidos simplemente en idea...” (T 2.3.9.2 SB 438; en el contexto de esta cita se identifica al bien con el placer y al mal con el dolor. Para Hume el bien es equivalente al placer y el mal al dolor, ver también: T 2.1.1.4 SB 276, 3.1.1.1 SB 399, 2.3.9.8 SB 439). La importancia de esta digresión sobre el placer y el dolor, en el ámbito de la discusión sobre la creencia, proviene de la relación que guardan con la motivación y la acción. Mi interés es examinar cómo la creencia, de acuerdo con Hume, contribuye también a la motivación y a la acción, de manera convergente con el placer y el dolor. Creo que esta convergencia es posible, desde un punto de vista humeano, en virtud de que la creencia misma es un tipo de sentimiento, sin menoscabo de sus propiedades cognitivas –i.e., su capacidad para aportar información relevante sobre una situación dada y formular los juicios correspondientes. Según explico en seguida, con estas consideraciones Hume prepara el camino para abordar uno de los problemas más interesantes relativos a la teoría de la acción.

Este problema consiste en determinar cuáles son las dificultades más fundamentales que una teoría de la acción de corte empirista tendría que resolver. Hume establece del siguiente modo este problema:

Si solo las impresiones influyeran sobre la voluntad, estaríamos sometidos en cada instante de nuestra vida a las mayores calamidades; en efecto, aunque podamos prever su aproximación, la naturaleza no nos ha provisto de un principio de acción que nos mueva a evitarlas. Y si, por el contrario, toda idea tuviera influencia sobre nuestras acciones, no sería nuestra condición mucho mejor, porque es tal la inestabilidad y actividad del pensamiento que las imágenes de cada cosa, especialmente bienes y males, están siempre cambiando en la mente; y si nos viéramos movidos por todas estas vanas concepciones no disfrutaríamos de un momento de paz y tranquilidad (T 1.3.10.2 SB 119)

El problema consiste en que si los humanos fuéramos un ente que solo reaccionara ante las impresiones sensibles seríamos incapaces de actuar, de elegir o de tomar decisiones. Esto muestra la dificultad que hay para explicar la relación entre las impresiones sensibles y la voluntad. Hume parece decir que la percepción de una impresión sensible no contiene por sí misma ningún principio de acción, en otras palabras, la percepción de que un coche se aproxima velozmente a nosotros por sí misma no contiene ningún elemento que nos ponga fuera de movimiento fuera de su camino. Ninguna impresión sensible, según esta idea, posee ninguna característica que pueda despertar en la voluntad un *motivo* para actuar de un modo o de otro. De este modo, nos sería indiferente la proximidad de un peligro o incluso de cualquier recompensa, porque para nosotros la experiencia del mundo sería *motivacionalmente* inerte. Por otra parte, si solo reaccionáramos ante las ideas, ante las impresiones provenientes de la memoria o de la imaginación, tampoco la pasaríamos bien. Según Hume, la inconstancia de nuestras ideas, su constante cambio, su falta de persistencia nos expondría a tales variaciones que difícilmente podríamos realizar proyecto alguno, así sea tan sencillo como preparar un alimento o atarnos las agujetas de los zapatos. Posiblemente, constantemente nos asaltarían nuevos proyectos y ante la percepción de nuevas ideas abandonaríamos pronto lo que hacemos. Pasaríamos de una idea a otra y la posibilidad de que nuestra voluntad pueda completar un proyecto iniciado en un momento dado parece mínima. Tal vez empezaríamos a preparar el desayuno, pero entonces el recuerdo de una canción nos brinca a la mente y corremos a buscarla en nuestra colección, pero mientras se nos antoja una bebida fría y esto ahora nos arrastra al bar más cercano, etc. También estaríamos expuestos a grandes peligros porque no reaccionaríamos adecuadamente ante los estímulos externos, pues sólo responderíamos a los estímulos provenientes de nuestra propia mente, tal como aparecen a nuestra atención en cualquier momento. Ante estas dificultades, una explicación adecuada sobre la acción humana parece requerir tanto de un modo de responder cómo las impresiones sensibles se conectan adecuadamente con la voluntad como una forma de explicar cómo ciertas ideas o percepciones logran mantenerse presentes ante la mente durante mucho más tiempo que su mera ocurrencia.

Creo que este requisito es tan fuerte que si no se satisface no hay posibilidad de contar con una buena explicación sobre la acción humana.

La respuesta de Hume a esta dificultad reside en la combinación de dos tipos de elementos: por una parte la creencia (por sus rasgos de solidez, firmeza y consistencia) y por otra parte el placer y el dolor que mueven todas nuestras acciones y pensamientos: “En la mente humana hay impresa una percepción del dolor y el placer, resorte capital y principio motor de todas sus acciones” (T 1.3.10.2 SB 118; T 3.3.1.2 SB 574).¹² Hemos visto que el placer o el dolor están asociados, de alguna manera, a cada percepción de la mente, así sea una percepción sensible o una idea. Además, tal percepción muchas veces sólo consiste en una expectativa de tal placer o dolor, la cual aún posee una gran fuerza motivadora (volveré más adelante, hablando de las pasiones y la motivación, a esta idea, v. p. 83). Ahora bien, citar tal expectativa, sin decir algo más, no parece suficiente para explicar cómo es que logramos satisfacer con éxito los proyectos motivados por ellas. Aun cuando supusiéramos que estas expectativas son de una duración tal que concluyen con la satisfacción misma del proyecto que motivan, aun así tendríamos que explicar cómo intervienen diversos estados intermedios e instrumentales para lograr realizar tales proyectos. Estos estados intermedios incluyen a la creencia y otras determinaciones de la mente relacionadas con la creencia. Hume lo explica del siguiente modo, en un pasaje que me parece contribuye a una explicación muy general sobre las condiciones fundamentales de la acción (y no sólo la acción humana):

Sabemos que al adaptar medios a fines somos guiados por nuestra razón e intención [*design*], y también que no es de una manera ignorante o casual como ejecutamos las acciones que tienden a la propia preservación, a obtener placer y evitar el dolor. Por consiguiente, cuando vemos en millones de casos que otras criaturas ejecutan acciones similares y las dirigen a fines similares, todos nuestros principios de razón y probabilidad nos llevan con fuerza invencible a creer en la existencia de una causa similar. En mi opinión, es innecesario ilustrar este argumento enumerando casos particulares. La más mínima atención nos dará más de los que pudieran ser necesarios. La semejanza entre las acciones de los animales y las de los hombres es tan perfecta a este respecto, que la primera acción

¹² El primer pasaje citado se inscribe específicamente en el contexto de discusión «la influencia de la creencia» sobre las pasiones y la acción y sobre él centraré mi análisis. El contexto del segundo pasaje abarca el examen del origen de las virtudes y vicios naturales. Hume relaciona fuertemente, de manera causal, al placer y dolor con las pasiones y la acción. Cito extensamente su pensamiento –al que volveré más adelante– porque Hume reúne ahí los elementos centrales de su teoría de la motivación o teoría de la acción (cf. Fate Norton 2000, 192). Teniendo presente estas consideraciones de Hume sostengo que la forma correcta de entender a las creencias es por su relación con las pasiones y la acción: “...cuando estas sensaciones [el placer y el dolor] desaparecen de nuestro pensamiento y sensibilidad, somos casi totalmente incapaces de experimentar pasión o acción, deseo o volición. Los efectos más inmediatos del placer y el dolor son los movimientos mentales de acercamiento y evitación, que se presentan como volición, deseo y aversión, tristeza y alegría, esperanza y miedo, según que el placer o el dolor cambien de situación y resulten probables o improbables, ciertos o inciertos, o sean considerados como estando fuera de nuestro poder por el momento” (T 3.3.1.2 SB 574).

del primer animal que queramos escoger nos proporcionará un argumento incontestable a favor de la presente doctrina (T 1.3.16.2 SB 176).

De acuerdo con esto, la razón tiene un papel instrumental para asegurar la satisfacción de los placeres y evitar el dolor a través de ciertas creencias. Su función en la explicación sobre la acción es la servir a la voluntad para mediar, por ponerlo así, entre las impresiones sensibles y las ideas, de manera que la voluntad pueda alcanzar los fines que se propone. Hume supone que los mismos elementos operan tanto en animales como en humanos. Veremos más adelante, cuando discutamos las creencias causales, que incluso la «razón y la probabilidad» a las que se refiere Hume aquí son producto o se explican por medio de las determinaciones naturales de asociación entre ideas (v. *infra* 1.5.3). El pasaje sugiere mucho más, pero si nos concentramos en la condición humana al respecto podemos examinar cómo la creencia influye sobre la voluntad.

Hemos visto que la creencia es una idea que ha ganado en fuerza y vivacidad, o bien, una idea que se ha vuelto más firme, estable y consistente que otras ideas. Mediante esta transformación de su fuerza fenoménica, la creencia mantiene en la mente por tiempo suficiente e intensidad suficientes la percepción de una cierta idea. Simultáneamente, la relación con el placer o el dolor logran el efecto motivacional adecuado, esto es, la capacidad para influir sobre la voluntad de tal manera que se satisface el proyecto de alcanzar cierto placer o rechazar cierto dolor. Hume explica este proceso en T 1.3.10.3 SB 119, del cual cito la idea central: “El efecto, pues, de la creencia consiste en levantar una idea simple hasta el mismo nivel de nuestras impresiones, y en conferirle una influencia análoga sobre las pasiones. Este efecto sólo puede tener lugar si se hace que una idea se parezca a una impresión en fuerza y vivacidad”. Este pensamiento, en resumen, es que la creencia aumenta la fuerza fenomenológica de una idea al nivel de las impresiones sensibles, de manera que la mente la percibe con la fuerza característica de éstas. Al mismo tiempo, esta fuerza logra influir o afectar a las pasiones de tal manera que las determinaciones naturales relativas al placer y al dolor se activan. Estas ideas de Hume logran explicar, al menos formalmente, las condiciones necesarias más básicas para explicar a la acción humana.

La acción humana, bajo esta perspectiva, reúne determinaciones cognitivas relativas a la experiencia, la imaginación y la asociación entre ideas. Por una parte, estas determinaciones explican parcialmente al estado de creer como un estado que eleva la intensidad fenoménica de ciertas ideas y, de este modo, hace que su percepción ante la mente sea «más sólida, firme y

consistente». Hume no parece ofrecer una prueba directa al respecto, pero sí una indirecta que proviene de la capacidad de la creencia para influir sobre la voluntad. Si esto es plausible, todavía tendríamos que explicar las condiciones que permiten superar la objeción consistente en que la mera percepción sensible no nos da motivos para actuar. En otras palabras, la percepción de objetos del mundo externo, por medio de impresiones sensibles, por sí misma no nos ofrece ningún motivo para actuar. De manera que se requiere de determinaciones de otro tipo para explicar cómo los agentes se involucran en proyectos diversos, así sea de largo plazo como terminar estudios de medicina o aun de corto plazo como hacer los preparativos para un viaje de fin de semana. De acuerdo con las ideas de Hume, estas determinaciones vienen de nuestra relación con el placer y el dolor y *fijan* nuestras preferencias, de esta manera logran producir *motivos* asociados con la experiencia sensible. El placer y el dolor explican que los objetos no nos sean indiferentes, así como que cultivemos preferencias específicas al respecto. Estas preferencias se ven apoyadas, o reflejadas, por las creencias. Creo que este es el contexto adecuado para entender la primera relación afirmativa que ofrece Hume en el *Tratado* entre las creencias y las pasiones: “Así como la creencia es un requisito casi indispensable para despertar nuestras pasiones, del mismo modo admiten éstas muy fácilmente la creencia; y por ello, no sólo los hechos que nos proporcionan emociones agradables sino también, y muy a menudo, los que causan dolor llegan a ser fácilmente objeto de fe y opinión” (T 1.3.10.4 SB 120). La relación entre creencias y pasiones es, de acuerdo con este modelo de la mente, muy estrecha y creo que es imprescindible explicarlas simultáneamente, en la medida de lo posible, para entender cabalmente la función distintiva de cada uno de los estados correspondientes. Por ahora, me parece claro que una manera correcta de entender la idea de que una creencia es una idea que ha ganado en fuerza se relaciona con las pasiones y otras emociones como una manera de señalar una de las condiciones *sine qua non* de la acción humana. Creo que en esto consiste la función práctica de la creencia. Por último, quiero terminar este capítulo examinando otra forma que adoptan las creencias del tipo (3): las creencias causales. Aprovecharé este examen para, simultáneamente, considerar el papel que la facultad de la razón tiene en la formación de creencias.

1.5 Las creencias causales y el papel de la razón en la formación de creencias

Dos consideraciones interrelacionadas invitan a considerar el papel de la razón en la formación de creencias causales. En primer lugar, el caso más interesante, para la teoría de la creencia de

Hume, es que estas creencias son un caso del tipo (3) de creencias sobre hechos. Parece haber dos casos generales en relación con las creencias causales, y cada uno se corresponde con los dos tipos de creencias sobre hechos. El primero es la percepción sensible *directa* de una relación entre objetos que puede describirse como causal. Por ejemplo, veo cómo tu mano se dirige al vaso y lo vuelca, derramando la cerveza sobre la mesa. Esta percepción da lugar a una creencia sobre hechos del tipo (2), por ejemplo: *creo que tu mano causó que se derramara la cerveza*. El segundo caso general consiste en una percepción sensible *parcialmente* directa de una presunta relación entre objetos que puedo describir como causal. Por ejemplo, ante la vista de una columna de humo proveniente del terreno baldío adyacente a mi calle surge en mi mente, espontáneamente, una creencia de la forma: *creo que un incendio causa al humo*. En este segundo caso, que representa un reto para Hume, tenemos que explicar cómo surge la creencia causal y cómo llegamos a la idea de que el humo fue *causado* por un incendio. Se trata de un caso de las creencias sobre hechos del tipo (3), esto es, es una creencia referida a algo de lo que no tenemos una impresión sensible directa, en este caso, el incendio. La peculiaridad, ahora, es que se trata de una creencia causal. Hume supone que este tipo de creencias son producidas por una «inferencia» (si bien emplea el término en un sentido amplio, más parecido a *derivación* que a operación lógica) mediante la cual la mente relaciona a una posible causa con un posible efecto. Surge aquí la primera consideración sobre la razón, porque ésta, en tanto que es una facultad que opera sobre inferencias, podría intervenir para realizar esta operación de derivación. En efecto, aunque desde el punto de vista del agente el paso de la percepción del humo a la formación de la creencia de que hay un incendio que lo causó parece inmediato y espontáneo, en realidad podría tener lugar un vertiginoso proceso racional inferencial. Además, la razón podría intervenir de otro modo, en la consideración misma de los supuestos racionales de la relación de causalidad.

Respecto de estos supuestos Hume es un escéptico de la capacidad humana de conocer satisfactoriamente a las relaciones causales, aunque no parece un escéptico de la causalidad misma. Por ejemplo, admite que “considerando el asunto *a priori*, cualquier cosa puede producir cualquier otra” e inmediatamente apunta al objetivo de su escepticismo: “...nunca descubriremos una razón por la que un objeto pueda ser o no causa de otro, por grande o pequeña que pueda ser la relación entre ellos” (T 1.4.5.30 SB 247). Así, dos objetos cualesquiera, cercanos o remotos, comparables o no, podrían relacionarse causalmente, pero en todo caso no podemos confiar en que podamos conocer adecuadamente dicha relación (cf. G. Strawson 2003). La razón más

importante para descartar que podamos acceder a tal conocimiento parece ser el carácter eminentemente *contingente* de la realidad y, por lo tanto, de las conexiones causales. En efecto, no hay ninguna *necesidad* en la relación causal entre dos objetos. Si la hubiera, entonces la razón podría conocerla, como conoce la necesidad en los enunciados que son sus objetos de conocimiento, como los enunciados de la aritmética y la geometría. En general Hume cree que la razón opera correctamente buscando conocimiento por medio de las relaciones entre ideas a las que llama “relaciones filosóficas”, por distinción con las relaciones naturales que hemos estudiado arriba.

Las relaciones filosóficas entre ideas se distinguen por ser relaciones o *comparaciones* entre ideas que establecemos deliberadamente. La preocupación filosófica por comparar objetos, dice Hume, tiene el sentido de “indicar cualquier asunto determinado de comparación, sin un principio de enlace” (T 1.1.5.1 SB 14). El «filósofo», como puede prescindir por completo de cualquier ‘principio de enlace’, puede comparar cualquier objeto con cualquier otro, sin necesidad de tener antecedentes o elementos comunes. A pesar de su evidente diversidad, Hume cree que este tipo de relaciones pueden reducirse a las siguientes siete: 1. Semejanza, 2. Identidad, 3. Espacio y tiempo (que incluyen las relaciones de distante, contiguo, arriba, abajo, antes, después, etc.), 4. Cantidad o número, 5. Cualidad, 6. Contrariedad y 7. Causa y efecto (cf. T 1.1.5.1-10 SB 14-15). En efecto, podemos comparar cualquier par de ideas de objetos por su semejanza (¿en qué se parecen, digamos, *El Quijote* y *Hamlet*?), por su identidad (¿cómo cambia el rosal que riego cada mañana?), por sus relaciones espacio-temporales (¿qué sucede allá enfrente?), por su cantidad (¿pesa más *El Quijote* o las pelotas de un partido de tenis?), por su cualidad (¿cuál es la naturaleza de las creencias?), por su contrariedad (¿de qué modo las pasiones se oponen a la razón?) y por su causa y efecto (¿qué provoca el humo en el terreno al final de mi calle?). Creencias importantes podrían surgir de estas relaciones. En particular, Hume sostiene que las verdades necesarias de la aritmética, del álgebra o la geometría surgen, precisamente, por relaciones filosóficas de este tipo. Las demostraciones formales serían operaciones en las que buscamos establecer la igualdad o proporción de los números, así como la identidad que hay entre números o magnitudes diversas y como se trata de relaciones entre ideas donde la mente procede con exactitud, podemos entonces estimar que estas son ciencias perfectas e infalibles (T 1.3.1.3-5 SB 70-71). Ahora podemos examinar las diferencias que hay entre las relaciones filosóficas y las relaciones naturales entre las ideas y apreciaremos una interesante asimetría, la cual es relevante para caracterizar a la

relación de causalidad. Este examen, además, nos permitirá señalar algunas diferencias entre las creencias del tipo (1), aquellas necesariamente verdaderas en virtud de alguna prueba deductiva y las creencias sobre hechos del tipo (2) y (3).

Respecto de las proposiciones que son verdaderas “por intuición o por demostración”, es decir, las verdades necesarias, como las proposiciones geométricas, aritméticas o lógicas, Hume sostiene que la propia prueba o comprensión de ellas conduce *inevitablemente* a la creencia o asentimiento correspondiente:

Con respecto a proposiciones que son probadas por intuición o demostración, la respuesta es fácil [la pregunta es: ¿en qué consiste la diferencia entre creer una proposición y no creerla?]. En ese caso, la persona que asiente no solamente concibe las ideas correspondientes a la proposición, sino que se ve *necesariamente determinada* a concebirlas de ese modo particular, sea directamente o por interposición de otras ideas. Todo lo absurdo es inteligible, y a la imaginación le es imposible concebir algo que se opone a una demostración (T 1.3.7.5 SB 95, énfasis añadido).

No parece, en efecto, que sea posible desarrollar, seguir y entender correctamente una prueba lógica –como en la aritmética o en la geometría– sin que necesariamente asintamos y creamos en la conclusión de tal prueba. Así, las verdades necesarias «son afirmadas de una manera diferente a como afirmamos las cuestiones de hecho» (cf. E 1.4.1.2. SB 25). Desde luego, las verdades sobre cuestiones de hecho están sujetas a contingencias y por ello no son necesariamente verdaderas ni necesariamente falsas (cf. T 1.3.6.5 89 y E 25). Ahora bien, la diferencia en el modo de asentir a unas y otras es compleja y reúne tanto a las propiedades distintivas de las proposiciones como a la fenomenología de los estados mentales relativos al creer. Al establecer relaciones filosóficas aspiramos a la exactitud y emprendemos una búsqueda deliberada de ciertas propiedades formales de las ideas (como la semejanza, la identidad, la proporción, el número, etc.), pero aun así el carácter necesario de estas propiedades formales, como vimos antes, conduce a los agentes a *aceptar necesariamente* ciertas proposiciones –so pena de irracionalidad o estupidez. Creer en una proposición formal es entender de algún modo el procedimiento inferencial o la demostración que la sustenta y los agentes no pueden rechazar el conocimiento que viene con tales procedimientos o pruebas. La asimetría mencionada consiste en que en las relaciones naturales entre ideas nuestra mente arriba de manera *espontánea* a ciertas ideas y aun a creencias y, en virtud de que estos productos espontáneos están sujetos a diversas condiciones contingentes que los pueden derrotar o privar de la certeza, los agentes tienen algún margen de decisión respecto de qué relaciones entre ideas aceptar o qué proposiciones creer. Mientras que mediante las relaciones filosóficas arribamos *necesariamente* a ciertas creencias, a pesar de que

se trata de relaciones buscadas deliberadamente. En las relaciones naturales, producidas espontáneamente por la mente, no asentimos necesariamente a las ideas o creencias que surgen. Los agentes, al respecto, pueden adoptar una actitud escéptica sobre sus creencias sobre hechos. Por ejemplo, me dicen que el rambután sabe a mango, pero puedo no asentir a este pensamiento y esperar a tener oportunidad de probar uno antes de pronunciar un juicio al respecto. O al ver una columna de humo en el terreno al otro lado de la calle pienso espontáneamente que hay un incendio, pero puedo no asentir inmediatamente a esta creencia y decidir reunir más elementos antes de conceder mi asentimiento (cf. Morris, 82).¹³ Es interesante observar que, adoptando esta actitud de cautela ante las ideas espontáneas, no podemos evitar arribar a otras concepciones o pensamientos (tal vez alguien hizo una fogata, están quemando llantas para una posada, etc.). Una vez que el agente adopta esta actitud escéptica, las propias determinaciones naturales de la mente que relacionan ideas hacen surgir en la mente nuevas relaciones entre ideas y el agente puede considerar entonces diversas posibles creencias a las que conceder su asentimiento. Al considerar estas posibilidades podemos apreciar qué es lo que hace que, para Hume, la causalidad sea una relación entre ideas que es tanto una relación natural como filosófica, *i.e.*, una manera tanto espontánea como deliberada de relacionar ideas entre sí. De esto se pueden derivar dos temas relevantes para el estudio de las determinaciones que dan lugar a las creencias.

Primero, los agentes disponen de un cierto margen de decisión para conceder su asentimiento a las creencias formadas por las relaciones naturales, pues al formarse espontáneamente por las determinaciones naturales de la asociación entre ideas ninguna necesidad conceptual o lógica compele a los agentes. El segundo tema consiste en determinar hasta qué punto la facultad de la razón influye sobre ese proceso de decidirse a asentir a una cierta idea. Este segundo tema es *normativo* y se refiere a la cuestión de hasta qué punto *deberíamos* creer en algo. Como hemos visto, según Hume creer es un acto espontáneo de nuestra naturaleza, tan natural como respirar, sin embargo, ahora estamos examinando cómo la razón podría, de alguna manera, *dirigir* la actitud escéptica de reservar o restringir nuestro asentimiento a ciertas consideraciones, esto es, la razón

¹³ Morris ha enfatizado este aspecto de la teoría de la creencia de Hume que, hasta donde puedo señalar, no ha sido apreciado por otros autores. La idea de Morris es que los agentes somos activos y tomamos ciertas decisiones sobre qué creer en relación con cuestiones de hecho. En mi traducción, dice Morris: “El énfasis de Hume en cómo *tomamos* el trabajo significa que tenemos un papel más activo en la formación de creencias de lo que sugieren sus explicaciones iniciales. No sólo tengo que tomar una decisión sobre cómo *Petra* [la antigua ciudad en Jordania] fue planeada, también tengo que considerar ‘las características y motivos’ de sus autores. [...] Muchas veces tomamos estas decisiones tácitamente, hacerlo de otro modo requiere que aceptemos evidencia testimonial adicional”, Morris, 82.

podría restringir el estado de creer. El estudio de las relaciones causales entre ideas nos permitirá examinar estas condiciones normativas en torno al estado de creer. Al respecto, parece haber dos dimensiones epistémicas: por una parte, las creencias causales surgen de la experiencia, de la costumbre y el hábito y son así ejemplos de relaciones *naturales* entre ideas. La razón no interviene en este proceso y, por ello, Hume adopta un punto de vista *negativo* sobre el papel de la razón para darnos conocimiento causal del mundo natural. Por otra parte, la razón parece sostener la actitud de refrenar nuestro asentimiento, esto es, permite examinar el grado de asentimiento que podemos conceder a tales o cuales consideraciones sobre hechos, incluyendo a las relaciones causales, antes de aceptar creencias al respecto. Así, la razón tiene un papel *positivo*, normativo, sobre el conocimiento relativo a las relaciones causales (cf. Kemp Smith, 382-383; Stroud, 92 y ss; Morris, 79). Me parece que este doble papel de la razón se relaciona, desde el punto de vista de Hume, de manera muy estrecha con las determinaciones naturales que producen a las creencias. Creo que una explicación completa sobre las creencias causales y empíricas, *desde un punto de vista humeano*, necesariamente tiene que reunir tanto a la capacidad normativa de la razón como a la experiencia como fuente de certeza y convicción. Kemp Smith ha tocado este punto con la siguiente reflexión:

[Dado que no todas las regularidades son confiables, que no todas las costumbres son buenas o que valen solo para un campo muy limitado de la experiencia, o que se deben a un cierto tipo de educación] Es por esto que, desde el punto de vista de Hume, el pensamiento reflexivo y los instrumentos por los cuales conduce la investigación científica y la organización social son, en un sentido, ‘artificiales’, pero a la vez son muy naturales e indispensables para el propósito de la vida humana (Kemp Smith, 383).

Desarrollaré en las últimas secciones de este capítulo estas ideas.

1.5.1 El punto de vista negativo: la razón no interviene en la formación de creencias causales

La razón debe exclusivamente sus inferencias confiables a “la comparación de ideas y al descubrimiento de relaciones que, mientras las ideas sigan siendo las mismas, sean inalterables” (T 1.3.3.2 SB 79). Sin embargo, esta confianza (o certeza) sólo es característica de las relaciones filosóficas entre ideas que dan lugar al conocimiento lógico y matemático, a saber, la semejanza, la proporción, la cualidad y contrariedad. Ninguna de estas relaciones permite explicar a la relación de causalidad –según Hume, dicha relación está constituida por la contigüidad, la

prioridad temporal y la conexión necesaria (cf. T 1.3.2.4-11 SB 75-77).¹⁴ Hume examina conceptualmente a la causalidad y rechaza que «la conexión necesaria» sea realmente constitutiva de la causalidad. Este examen conceptual mismo es una operación filosófica, dirigida por la razón, y es un ejemplo de cómo la razón *positivamente* establece criterios normativos para descartar que *ella misma* pueda producir creencias causales. Al mismo tiempo, indica cuáles son los elementos que sí constituyen a la causalidad y, en consecuencia, bajo qué condiciones podríamos conceder asentimiento a relaciones causales. Revisaré a continuación las líneas generales de este examen conceptual.

Para empezar es necesario que exista una cierta *contigüidad* entre la causa y el efecto para que tenga lugar alguna relación entre ellos. Ésta puede ser mediata o inmediata y, *por lo menos*, consistir en una «una cadena causal que conecta a dos objetos remotos» (cf. T 1.3.2.6 SB 75).¹⁵ La prioridad de la causa sobre el efecto indica que la relación causal es temporal: la causa es anterior al efecto. Finalmente Hume argumenta que la idea de causalidad requiere que se pueda pensar la relación entre la causa y el efecto de manera que no pueda concebirse a la una sin el otro: “Hace falta una CONEXIÓN NECESARIA” (T 1.3.2.11 SB 77, vérsales de Hume). En efecto, dos objetos podrían ser concomitantes e incluso uno presentarse con anterioridad al otro sin estar relacionados causalmente. En estos casos observamos tanto la contigüidad, inmediata o mediata entre dos objetos, como la prioridad de uno ante el otro y aun así podría no existir ninguna relación causal entre ambos, por ejemplo: Mi gato brinca sorpresivamente sobre la mesa y enseguida suena el teléfono, pero no hay ninguna relación causal entre ambos hechos. Con base en esto, Hume sostiene que la causalidad requiere de una cierta conexión entre la causa y el efecto, la cual es «mucho más importante» que la contigüidad y la prioridad. Esta idea es la que parece explicar mi creencia en que un incendio provocó la columna de humo que veo junto a mi calle. Mi creencia

¹⁴ La argumentación de Hume es más ambiciosa de lo que estoy presentando. Hume dirige su examen sobre la tesis metafísica: “*todo lo que empieza a existir debe tener una causa de su existencia*” (T 1.3.3.1 SB 78). Su conclusión negativa sobre esta tesis es lapidaria: “...mediante un examen adecuado encontraremos que toda demostración que haya sido presentada en favor de la necesidad de una causa es falaz y sofisticada” (T 1.3.3.4 SB 80). El argumento que presento en el texto conduce a esta afirmación, pero enfatizaré las limitaciones de la razón para explicar adecuadamente a la causalidad.

¹⁵ Stroud ha impugnado que la ‘contigüidad’ sea realmente esencial a la causalidad. Observa que Hume no está buscando exactamente las notas esenciales de la causalidad, sino las impresiones de las que derivamos tal idea. Así que es evidente que “no tenemos una impresión de contigüidad *cada vez* que observamos un par de objetos que consideremos relacionados como causa y efecto”. Con este antecedente plantea su objeción: “¿Cómo sabemos entonces, si es que podemos saberlo en absoluto, que la contigüidad es “esencial” a la causalidad?” Observa, además, que en el *Ensayo* Hume no menciona a la contigüidad como esencial a la causalidad. Cf. Stroud, 68-69.

descansa en que pienso que la relación o conexión entre el humo y el fuego es *necesaria*, en el sentido de que estos objetos parecen inseparables entre sí. La creencia refleja esta fuerte relación y parece confirmar la importancia de la conexión necesaria para formular creencias causales. La creencia misma no parece otra cosa que la convicción en la conexión necesaria entre esos objetos. Con todo, esta relación es cuestionable.

No es inconcebible, ni lógica ni empíricamente, separar ambos objetos, en este caso, el fuego y el humo, de manera que podamos concebirlos aisladamente. Así, aun cuando la necesidad parece calificar a nuestra certeza en las conexiones causales entre objetos, la razón rápidamente descubre que se trata de una relación dudosa. En consecuencia, la creencia al respecto es dudosa. Sin embargo, como he dicho, la creencia causal no parece otra cosa que el reconocimiento de la relación causal necesaria entre dos objetos. Lo que nos ha enseñado el análisis conceptual de la causalidad es, en efecto, la idea de que sin la necesidad no parece que podamos explicar a las creencias causales. Esta certeza constituye a las creencias causales y, aunque se descubra ahora como dudosa, Hume tiene que explicar cuál es su origen, es decir, cómo llegamos a adquirir esta certeza. Desde la perspectiva de Hume, esta dificultad proviene de que para él una idea significativa tiene que originarse en una impresión o impresiones primarias, de las que tal idea “puede haberse derivado” (cf. T 1.3.2.12 SB 77) —es decir, «el principio de la copia», como hemos visto, que implica que toda idea sea una copia de alguna impresión sensible anterior a ella. Así, no encuentra, porque no hay tal cosa, ninguna impresión primaria que de origen a la idea de ‘conexión necesaria’ —en cambio, sí las hay de las ideas de ‘contigüidad’ directa y de ‘prioridad’. Esto torna aún más difícil explicar qué queremos decir con ‘conexión necesaria’. La dificultad se acrecienta al observar que *es un hecho que contamos con tal idea*, así que si no se origina en una impresión primaria, sigue siendo pertinente preguntar cómo surge en nuestra mente tal pensamiento.

Como hemos dicho arriba, la causalidad no es objeto del conocimiento formal propio de la lógica y de la matemática, por lo tanto, no es posible para la razón establecer ninguna relación necesaria entre la causa y el efecto. En otras palabras, la idea de un objeto no contiene ninguna idea de qué efectos puede producir (o bien, que puede producir *cualquier* efecto, lo que tampoco nos permite discernir nada positivamente sobre efectos particulares) ni tampoco contiene la idea de que sea el efecto de tal o cual otro objeto particular. Esta idea posiblemente sea controversial: al ver a un conejo particular sé que fue *producido* por otro conejo. Es decir, la idea del conejo

parece *contener* que es el efecto de algún otro conejo anterior –aunque tal vez no pueda decir exactamente cuál conejo. Sin embargo, el punto de Hume parece ser que no podemos, racionalmente, aceptar que ésta es una idea *necesaria*. Diversas contingencias podrían explicar al conejo actual y no parece implicar ninguna contradicción pensar que tal o cual conejo es el efecto de *ningún* otro conejo particular. Esto parece un disparate en términos naturales, pero no lo es en términos lógicos. Hume parece tener un pensamiento de este tipo cuando escribe: “Si alguien dijera que es sólo probable que el sol salga mañana, o que todos los hombres deben morir, haría el ridículo. Y sin embargo, es evidente que no tenemos más seguridad en estos hechos que la proporcionada por la experiencia” (T 1.3.11.2 SB 124). De modo que la idea de cualquier objeto no contiene ni la idea de los efectos que lo han producido ni de las causas que puede producir. Esta razón es central en el siguiente argumento:

...como todas las ideas distintas son separables entre sí, y las ideas de causa y efecto son evidentemente distintas, nos resulta fácil concebir cualquier objeto como no existente en este momento y existente en el siguiente, sin unirle la idea distinta de causa o principio productivo. ...por consiguiente, es de tal modo posible la separación real de estos objetos, que ello no implica contradicción ni absurdo alguno, por lo que dicha separación no puede ser refutada por ningún razonamiento efectuado en base a meras ideas; y sin esto es imposible demostrar la necesidad de una causa (T 1.3.3.4 SB 79-80).

Con este argumento Hume descarta que la idea de conexión necesaria pueda ser comprensible como una inferencia de la razón. La razón no puede establecer la necesidad lógica que vincularía en ningún caso a la causa con el efecto. Pero esto no parece suficiente para minar la confianza que tenemos en las creencias causales. Es decir, a pesar de que no contamos con una explicación satisfactoria de la idea de conexión necesaria, ni como idea asociada o derivada de una impresión sensible original ni como idea derivada inferencialmente de otras ideas o impresiones, aun así tenemos confianza en las creencias causales. Por ejemplo, la creencia que relaciona causalmente al fuego con el humo nos parece, a primera vista, necesaria, pero aun cuando esta necesidad pueda cuestionarse en todo caso, nos parece una creencia de primera importancia práctica. La teoría de la creencia de Hume todavía nos debe una explicación al respecto.

1.5.2 El punto de vista positivo: la costumbre, o la experiencia, explican a las creencias causales

Sabemos ahora que este tipo de creencias no puede originarse en la mera consideración racional de las ideas. Estrictamente, el argumento anterior ha mostrado que las ideas de fuego y humo están separadas «en la imaginación» (T 1.3.3.3 SB 80) en el sentido de que podemos pensar, *sin*

contradicción, al uno sin el otro, y viceversa.¹⁶ Con todo, aun cuando no podamos explicar cómo, no parece haber duda de que la idea de la causalidad se origina en la experiencia (cf. T 1.3.1.1 SB 69). De acuerdo con el modelo de la mente de Hume, tal idea tendría que originarse, de alguna manera, por *derivación* de ciertas “impresiones o, al menos, con ideas de la memoria, que equivalen a impresiones” (T 1.3.4.1 SB 82). En este sentido de inferencia, como derivación, Hume parece establecer las condiciones formales de la idea de causalidad:

...cuando hacemos una inferencia desde las causas a los efectos tenemos que establecer la existencia de estas causas. Y sólo tenemos dos formas de hacer esto: o por una inmediata percepción de nuestra memoria o sentidos, o por inferencia de otras causas. Y de estas causas deberemos cerciorarnos del mismo modo: o por una impresión presente, o por inferencia de las causas de *estas otras* causas, y así sucesivamente, hasta llegar a algún objeto que veamos o recordemos. Nos es imposible proseguir nuestras inferencias *in infinitum*; y lo único que puede ponerlas término es una impresión de la memoria o de los sentidos... (T 1.3.4.1 SB 83)

De este modo, la creencia sobre la causa -o el efecto- que pretendemos inferir se deriva de una impresión presente y, en último término, de la experiencia misma. Como hemos visto, por medio de la relación con una impresión sensible y en virtud del principio de transmisión de fuerza de la impresión la creencia que pretendemos inferir obtendrá la fuerza que la distinguirá de una mera idea o una mera concepción (dicho principio nos recuerda que la fuerza relativa de una impresión y una idea se transmite siempre de la impresión a la idea, cuando están relacionadas del modo en que ahora pretendemos explicar).¹⁷ Con esta observación Hume logra explicar cómo la inferencia causal que buscamos se hace plausible como una fuente de creencias confiables: “No es sino la fuerza y vivacidad de la percepción lo que constituye el primer acto del juicio y pone las bases de ese razonamiento, construido sobre él, cuando inferimos la relación de causa y efecto” (T 1.3.5.7 SB 86). Llegamos nuevamente, en un contexto distinto, a la tesis central de esta teoría de la creencia: *Las creencias sobre cuestiones de hecho se caracterizan por su relación con impresiones de los sentidos o de la memoria* (cf. T 1.3.5.7 SB 86). Con base en estos antecedentes, Hume explica ahora a la certeza en las creencias causales. La discusión anterior le permite descartar a la necesidad como un componente conceptual constitutivo de la idea de causalidad, pero a la luz de

¹⁶ Así como la imaginación permite asociar ideas, bajo el principio de asociación, del mismo modo puede descomponerlas. De este modo, nos permite «visualizar» al fuego sin humo. La imaginación, al parecer, es para Hume la facultad que explica al proceso de abstracción.

¹⁷ Con base en esto, Hume explica cómo obtenemos confianza en las creencias históricas: son derivaciones causales que se originan, en algún momento, en alguna impresión: “...creemos que César fue asesinado en el Senado durante los *idus de marzo*. Y creemos esto por hallarse establecido por el testimonio unánime de los historiadores... Estas ideas estaban o en la mente de quienes se hallaban inmediatamente presentes en esa acción y recibieron directamente las ideas de su existencia o han sido derivadas del testimonio de su existencia” (T 1.3.4.2 SB 83).

la relación de la causalidad con la experiencia, encuentra que existe *al menos* una ‘conjunción constante’ entre lo que identificamos como causa y lo que identificamos como efecto. Por ejemplo, la experiencia de la relación entre el fuego y el humo consiste en el recuerdo compuesto de haber experimentado constantemente relacionados al fuego y el humo. Este recuerdo, por sí mismo, consiste en la conjunción constante de la impresión de cada uno de estos objetos. La conjunción constante parece caracterizar a la relación causal:

Y de la misma manera recordamos mentalmente su conjunción constante en todos los casos pasados. Sin más preámbulos, llamamos a los unos *causa* y a los otros *efecto*, e inferimos la existencia de unos de la de otros. En todos los casos por los que sabemos de una conjunción entre determinadas causas y efectos, tanto estas causas como estos efectos han sido percibidos por los sentidos y están presentes a la memoria (T 1.3.6.2 SB 87).

Así, desde el punto de vista de la certeza de los agentes en las creencias causales, la conjunción constante reemplaza a la conexión necesaria, y ésta es una tesis normativa sancionada por la argumentación racionalidad sobre la idea de causalidad. De este modo, “para poder decidir si dos objetos son causa y efecto”, la ‘conjunción constante’ es el elemento que complementara a la ‘contigüidad’ y ‘sucesión’ y, adicionalmente, ocupará el lugar de la conexión necesaria como elemento esencial a la relación de causa y efecto (cf. T 1.3.6.3 SB 87). En resumen, los elementos esenciales de la idea de causalidad son las nociones de ‘conjunción constante’, ‘contigüidad’ y ‘sucesión’ (prioridad). Negativamente, esta tesis, aunque propuesta mediante la razón, excluye que la razón tenga papel alguno en la formación de creencias causales. Hume así lo afirma, contundentemente: “cuando la mente pasa de la idea o impresión de un objeto a la idea de otro, o creencia en él, no está determinada por la razón, sino por ciertos principios que asocian las ideas de estos objetos y las unen en la imaginación” (T 1.3.6.12 SB 92). De este modo, su argumento es *negativo* respecto de que la razón pueda dar lugar a las creencias causales o sobre cuestiones de hecho y, al mismo tiempo, es *positivo* respecto de que es la imaginación la facultad que da lugar a tales creencias (cf. Stroud, 92 y ss.). Esta faceta positiva explica el hecho de que tenemos creencias fácticas causales. El argumento muestra los poderes relativos y distintivos de la razón y de la imaginación para explicar y producir este tipo de creencias. Además, y de manera muy importante, el argumento muestra la distinta relación de ambas facultades con la uniformidad de las percepciones de la mente, según veremos en seguida.

1.5.3 La discusión de la probabilidad como origen de las creencias causales

Si la conjunción constante es el elemento conceptual constitutivo de la idea de causalidad, cabe suponer que, con base en él, la razón pudiera realizar operaciones inferenciales de las que se derivaran creencias causales fácticas. En efecto, de acuerdo con Hume, la facultad de la razón realiza operaciones deductivas sobre ideas y creencias o sobre principios generales de los que infiere diferentes juicios o creencias.¹⁸ Bajo estas condiciones, respecto de las creencias causales la razón podría proceder “según el principio de que *casos de los que no hemos tenido experiencia deben ser semejantes a aquellos en que sí la hemos tenido, pues la naturaleza sigue siempre uniformemente el mismo curso*” (T 1.3.6.4 SB 89).¹⁹ Este principio, en las inferencias causales, opera como una premisa universal o general de la cual la razón *podría* inferir juicios particulares relativos a la causa o al efecto. Stroud llama a éste el «principio de uniformidad» (Stroud, 84). Ahora bien, de acuerdo con el argumento de Hume, si el principio fuera aceptable, entonces podríamos inferir confiadamente de cualquier percepción la idea correspondiente relativa a la causa o al efecto, de acuerdo con aquello que quisiéramos explicar. Por ejemplo, si el principio de uniformidad fuera correcto, entonces cabría suponer que la relación causal entre el fuego y el humo es constante o uniforme, de tal manera que no podríamos aceptar la percepción de uno sin la percepción del otro. Así, si percibo fuego o humo, y me interesa determinar la causa o el efecto correspondiente, podría inferirlos a partir del principio de uniformidad y la percepción particular. Antes de pretender avanzar más en la descripción de este procedimiento inferencial, Hume encuentra que hay buenas razones para dudar del principio de uniformidad y, por lo tanto, de la inferencia causal. El examen de estas razones introduce la consideración de que la probabilidad, o la inducción, sea el tipo de inferencia racional que produzca a las creencias causales.

En primer lugar, el principio parece falso: “[p]odemos al menos concebir un cambio en el curso de la naturaleza, y ello prueba suficientemente que tal cambio no es absolutamente

¹⁸ El álgebra y la aritmética, según Hume, son los campos donde la razón realiza plenamente sus facultades, mostrando un “criterio preciso [*a precise standard*] para juzgar de la igualdad y proporción de los números”, porque ellas “son las únicas ciencias en que puede efectuarse una argumentación de cualquier grado de complejidad, conservando sin embargo una exactitud y certezas perfectas”, v. T 1.3.1.5 SB 71.

¹⁹ Uno de los propósitos del *Tratado* es investigar los principios generales que rigen a la mente. Hume realiza esta tarea investigando el pensamiento y las relaciones que hay entre ideas (cf. T Intr. 10 SB XIX, 1.1.3.1 SB 8 y 1.1.4.6 SB 12-13 (donde introduce el principio de atracción entre ideas, el cual Hume consideraba su mayor descubrimiento). En EC sostiene que una investigación metafísica que no sea especulativa, sino rigurosa, consiste en examinar los principios que rigen la mente, regulan el entendimiento, estimulan los sentimientos, y nos hacen aprobar o rechazar a cualquier objeto particular, acción o conducta (EC 6 y ss).

imposible” (T 1.3.6.5 89; cf. E 25). El carácter contingente de los sucesos naturales, tal como los entendemos, impiden aceptar que el principio de uniformidad sea verdadero y no puede ofrecer, en la posible inferencia causal deductiva, el respaldo general necesario para las inferencias causales que buscamos. Hume, entonces, sostiene que «no puede haber ningún argumento demostrativo [o deductivo] para probar que aquellas casos, de los que no tenemos experiencia, serán como aquellos de los que sí hemos tenido» (cf. T 1.3.6.5 SB 89). En otras palabras, dado que el principio de uniformidad es falso, no podemos inferir deductivamente de una observación particular nada sobre sus posibles causas o efectos.²⁰ Hume concluye entonces que *demonstrativamente* (o deductivamente) no podemos inferir ni la causa ni el efecto de una impresión dada.²¹ Desde luego, esto no anula la veracidad de la observación respecto de la uniformidad del flujo de las ideas en la mente. Lo único que necesita el modelo de la mente de Hume para explicar cómo opera el principio de transmisión de fuerza de las impresiones a las ideas y la relativa regularidad de la asociación entre ideas es un cierto grado de uniformidad. Podemos observar, en efecto, que el flujo de nuestras ideas no siempre es tan regular o uniforme, o nos sorprendemos pensando en algo que, aparentemente, no tiene relación alguna con nuestra situación presente. Estas excepciones no son significativas respecto del grado de uniformidad que predomina en el flujo de las percepciones de la mente. Con todo, la prueba deductiva no es el único recurso de la razón. Las pruebas de «probabilidad» –hoy diríamos inductivas- son otro recurso de la razón que podría explicar a las creencias derivadas de inferencias causales.

Para Hume en las pruebas de probabilidad la razón no descubre meramente las relaciones entre las ideas, “consideradas en tanto que tales”, en vez de ello, descubre las relaciones que hay

²⁰ Hume ilustró claramente este punto sobre la incapacidad de nuestro conocimiento para determinar las causas efectivas de un suceso: “Hay causas que son absolutamente uniformes y constantes en la producción de determinado efecto, y jamás se ha encontrado fracaso o irregularidad alguna de su operación. Siempre ha quemado el fuego, y el agua siempre ha ahogado a cualquier criatura humana. [...] Pero hay otras causas que se han mostrado más irregulares e inciertas. El ruibarbo no resulta siempre un purgante ni el opio un soporífero para todo el que toma estas medicinas. Es cierto que, cuando cualquier causa deja de producir el efecto usual, los filósofos no atribuyen esto a una irregularidad de la naturaleza, sino que suponen algunas causas secretas, existentes en la singular estructura de los componentes, han impedido esta operación” (EC 57-8).

²¹ Este es un buen punto respecto de las limitaciones para establecer deductivamente relaciones causales, pero, como ha observado Stroud, no es un buen punto para descartar la validez de las inferencias deductivas (Stroud, 92). Hume parece tener razón en sostener que no podemos derivar relaciones causales sobre hechos contingentes por medio de la razón, pero esto no es suficiente para cuestionar la validez de su capacidad para hacer deducciones. Hume, me parece, no pretende decir otra cosa, p.e.: “...no pueden existir argumentos *demonstrativos* que prueben que *casos de los que no hemos tenido experiencia son semejantes a aquellos en los que sí la hemos tenido*” (T 89), que es una afirmación sobre los límites de los argumentos demostrativos (o deductivos) y no sobre la capacidad de la razón para operar deductivamente.

realmente entre los objetos, por lo que “tiene que basarse en algún respecto en las impresiones de nuestra memoria y sentidos, y en algún otro, en nuestras ideas” (T 1.3.6.6 SB 89). En estas pruebas la razón considera alguna impresión (que podría provenir de los sentidos o del *almacén* de la memoria) y alguna idea. De ambos elementos, *tendría que* derivar inductivamente la idea del elemento causal que nos interesa, ya sea la causa o el efecto. Si no se cumple esta condición, la única alternativa posible entonces sería que el resultado de la inferencia fuera otra sensación, pero la relación entre sensaciones se explica por alguna otra determinación *natural* de la mente, ya sea la contigüidad, la semejanza o la relación de causa y efecto. No olvidemos que aun buscamos una relación inferencial racional, *deliberada* («filosófica»), a partir de una sensación y una idea. Ahora la buscaremos bajo la condición señalada. Hume establece ahora otra condición, pero inmediatamente comienzan a apreciarse las dificultades para la razón.

Si la razón pretende explicar a las inferencias causales, entonces *tiene que* recurrir a la relación de causalidad: “La única conexión o relación de objetos que puede llevarnos más allá de las impresiones inmediatas de nuestra memoria y sentidos es la de causa y efecto” (T 1.3.6.7 SB 89). No puede ser de otro modo, para Hume, porque esta relación es la única que podría conducir a la mente de una idea a otra idea de un objeto a otro del que no tenemos una impresión directa. Además, la relación de causa y efecto ancla uno de sus términos en la experiencia. Como hemos descubierto, por experiencia sabemos “que ciertos objetos particulares han estado constantemente conectados entre sí en todos los casos pasados” (T 1.3.6.7 SB 90). Sin embargo, en este punto tenemos que volver a considerar al principio de uniformidad: como la razón *supone* que “un objeto similar a aquéllos está inmediatamente presente en la impresión, conjeturamos la existencia de un objeto similar a su habitual acompañante”. Por ejemplo, si veo una columna de humo cerca de mi casa, infiero la idea de que hay un incendio, porque en el pasado he tenido la *impresión conjunta* de humo y fuego. Por ello, asumiendo que este caso es comparable a los anteriores, infiero la idea de que hay un incendio, si bien no tengo impresión alguna de *este* incendio. Se satisfacen, aparentemente, todas las condiciones para aceptar a la inferencia probabilística de la razón, pero aun así cabe la siguiente objeción.

De acuerdo con Hume tal prueba es circular, porque depende del principio de uniformidad y éste, a su vez, depende de la relación de causa y efecto. En otras palabras, sabemos por experiencia que ciertas relaciones entre objetos son constantes y, con base en ello, se establece el principio de uniformidad. Una vez llegados a este punto, derivamos del principio de uniformidad que aquellos

casos de los que no tenemos experiencia son semejantes a los anteriores. Hume se refiere a esta circularidad en los siguientes términos: “El mismo principio no puede ser a la vez causa y efecto de otro” (T 1.3.6.7 SB 90). Explicando esta circularidad señala que “la probabilidad está basada en la conjetura de que existe semejanza entre objetos de los que hemos tenido experiencia y objetos no experimentados” (T 1.3.6.7 SB 90). La *semejanza*, en efecto, es la determinación de la mente que garantiza tanto a las relaciones causales como al principio de uniformidad. Como relación natural, la semejanza es una relación poderosa y constituye una de las determinaciones naturales más fuertes de la asociación de ideas. El principio de uniformidad surge de esta relación natural.²² Puede observarse ahora que el principio de uniformidad va mucho más allá de establecer semejanzas entre ideas actuales de la mente.

El principio de uniformidad permite extender la uniformidad de las experiencias pasadas hacia las futuras. Indica que cabe esperar que las experiencias futuras, anticipadas ya, sean semejantes a las pasadas. Esto es una suposición sobre el futuro y, desde la perspectiva de Hume, su origen constituye una de las determinaciones de la mente para extender sus conjeturas y relacionar de este modo ideas. Esta determinación hace surgir al tipo de creencias fácticas que consideramos y constituye una fuente de certidumbre o confianza que descansa en la experiencia pasada. Price, por ejemplo, ha considerado que este aspecto “constructivo” del *Tratado* permite sostener que en «la esfera de las cuestiones de hecho las creencias razonables o justificables o sentidas son aquellas –y sólo aquellas– que están basadas en la experiencia de estas regularidades» (Price, 7§6). Baier aprecia la misma confianza que da la experiencia de ciertas regularidades a las creencias sobre cuestiones de hecho. De acuerdo con ella, «el verdadero sentido del Libro III del *Tratado* [en donde Hume propone la teoría que presento ahora] es la positiva tesis de que la experiencia nos salva cuando la razón deductiva nos deja caer» (Baier 1991, 56). En efecto, hemos visto que el aspecto destructivo (o escéptico) del *Tratado* consiste en señalar que la razón es insuficiente para explicar las relaciones causales, mientras que el aspecto positivo recupera tales relaciones, así como las creencias fácticas derivadas de ellas, mediante la experiencia y la determinación de las regularidades que la constituyen. Me parece que podemos ilustrar este pensamiento con una consideración relativamente sencilla.

²² Desde luego, la semejanza también es una relación filosófica, pero en este caso su función establece relaciones entre objetos o entre ideas de los que ya tenemos alguna idea y de las que hemos detectado ciertos rasgos comparables. Como relación natural, precisamente, la semejanza relaciona espontáneamente ideas.

La causalidad, en tanto relación filosófica o deliberada entre dos ideas, es la respuesta a una cierta actitud inquisitiva en la que, ante una determinada experiencia sensible, queremos saber o la causa o el efecto de tal experiencia. Percibimos un cierto suceso y queremos saber cómo se relaciona causalmente con otros. Esta curiosidad despierta espontáneamente en nosotros y, me parece, se satisface de dos modos básicos. O nos detenemos a especular sobre las relaciones causales de la impresión ante nosotros o, sencillamente, examinamos *directamente* cuáles son tales relaciones. El caso interesante es el primero, porque en el segundo, si tenemos éxito, lograremos creencias explicativas asociadas directamente con ciertas percepciones. En el primer caso opera el principio de uniformidad y da lugar a creencias. Un ejemplo: escucho voces del otro lado de la habitación, pero sé que en la casa sólo estoy yo. Examinó las causas que pueden dar origen a tal impresión y encuentro que las voces se originan en que se ha activado automáticamente el radio-reloj despertador. La búsqueda de esta causa es deliberada. No tengo la impresión directa del despertador sonando, pero las voces, su origen, su *semejanza* con emisiones anteriores, mi conocimiento de las condiciones en mi casa, saber que estoy solo, etc., conducen a la idea de que el radio se prendió automáticamente. Deliberadamente busco la causa de este efecto y lo hago disponiendo de los materiales que me suministra la experiencia presente y pasada. La imaginación, eventualmente, arriba a la creencia de que las voces surgen del radio despertador. De acuerdo con otros marcos conceptuales y con mis sistemas generales de creencias (otra forma de presentar al principio de uniformidad), mi especulación sobre la causa de las voces seguirá las pautas uniformes que me son habituales. Con la memoria y la imaginación recorro los objetos en la habitación, trato de discernir las voces, su orientación, imagino su fuente, etc. Pronto llego a la idea de que se trata del radio activado automáticamente. Me siento satisfecho con esta explicación, *le concedo mi asentimiento* y llego a esta creencia: las voces provienen del radio-reloj despertador. Las creencias causales sobre cuestiones de hecho, como ésta, tienen su origen último en la experiencia y en las determinaciones naturales de la mente a las que ésta obedece. Son las propias determinaciones naturales las que hacen posible a la relación filosófica de la causalidad, la cual, por ello, se distingue de las otras relaciones filosóficas. Así, no es la razón la que opera la relación filosófica de la causalidad, sino la imaginación:

Aun ayudada por la experiencia y observación de su conjunción constante en todos los casos pasados, la razón no puede mostrarnos nunca la conexión de un objeto con otro. Por tanto, cuando la mente pasa de la idea o impresión de un objeto a la idea de otro, o creencia en él, no está determinada por la razón, sino por ciertos principios que asocian ideas de estos objetos y las unen en la imaginación (T 1.3.6.12 SB 92).

De modo que, aun cuando deliberadamente busquemos relaciones causales entre ideas relativas a cuestiones de hecho y encontremos relaciones correctas entre ideas, tendremos que admitir que tales procesos inferenciales requieren del principio de uniformidad y que dependen en último término de las determinaciones naturales de la mente. De este modo, esta teoría de la creencia nos presenta el papel cognitivo desempeñado por la creencia, pero la fuente de tales creencias es la imaginación o las determinaciones naturales de la mente.

2. Elementos causales y cognitivos de la teoría de las pasiones de Hume

Tanto su teoría de las ideas como su modelo de la mente le señalan a Hume cómo entender y explicar a las pasiones. De acuerdo con esto, las pasiones tendrían que ser alguna percepción en la mente, esto es, impresiones o ideas. Además, las mismas determinaciones básicas que, como hemos visto, relacionan causalmente a las impresiones y a las ideas, dan lugar a los diversos estados mentales. La operación conjunta de estos materiales y determinaciones explica también a las pasiones. En esta sección me concentraré en el estudio de estos elementos y en cómo su operación conjunta explica, según ideas de Hume, a las pasiones. Me concentraré en los procesos causales que dan lugar a las pasiones y mostraré cómo en estos procesos intervienen las determinaciones básicas de la mente, como el placer, el dolor, el deseo de belleza, la semejanza y la igualdad. Estas determinaciones logran las asociaciones naturales entre ideas que caracterizan a la vida mental. Posteriormente, mostraré cómo estas dos últimas determinaciones explican a la simpatía y haré un primer apunte de la importancia que tiene esta relación para la moralidad. Además, siguiendo ideas de Anette Baier (1990) y de Lili Alanen (2005) defiendo una interpretación cognitiva de las pasiones.

Esta interpretación descansa en dos ideas básicas. Primera, las pasiones son *acerca de algo* (cf. Baier 1990, 4), esto es, poseen algún contenido intencional, un objeto al que se refieren. Segunda, en la observación, según Baier (y que Alanen acepta) de que tal objeto se descompone en tres objetos: tendrían un objeto *profundo* que aflora con cada manifestación de la pasión. Un objeto *formal* que recoge los rasgos formales de la pasión y, por último, un objeto *inmediato*, esto es, aquel objeto percibido por el agente de manera inmediata y que desencadena los complejos mecanismos de la pasión. El objeto profundo subyace o se esconde detrás de cada objeto inmediato. La percepción del objeto inmediato suscita que se reaviven una especie de memoria emocional de este tipo de objeto, con el cual el agente se ha relacionado a lo largo de su vida. El objeto formal recoge las características distintivas o definitorias de la pasión, por ejemplo, los rasgos que hacen al orgullo ser lo que es, independientemente de los objetos inmediatos o profundos que lo suscitan en un agente dado. Desarrollaré estas ideas en la sección 2.2.1.

2.1. Las pasiones calmadas y el deseo de belleza

Según su modelo de la mente las pasiones son para Hume una forma de impresión: “incluyo bajo este nombre [impresiones] todas nuestras sensaciones, pasiones y emociones tal como hacen su primera aparición en el alma”. Lo que parece separar a las pasiones de las ideas es el «mayor grado de fuerza y vivacidad» de las primeras (cf. T 1.1.1.1 SB 1-2). Como hemos visto en el caso de las creencias, la noción de ‘fuerza’ de Hume tiene un sentido especial. Respecto de las pasiones este sentido parece referirse al modo en que un cierto estado mental *domina* a la mente por un periodo dado y de una manera distintiva. Este periodo es el tiempo que dura la pasión o el tiempo en que un sujeto la experimenta. Explicar la manera distintiva nos va a requerir más tiempo. Para empezar, parece componerse de dos aspectos generales los cuales no necesariamente se manifiestan ni operan juntos: primero, una cierta sensación fenomenológica bajo la forma general de placer o dolor y, segundo, una relación con la acción, lo que identificaré como su papel motivador de la acción (cf. Fate Norton 2000, I64-165; Karlsson, 263; Baillie, cap. 4).²³ Consideraré ahora cómo estos dos aspectos se relacionan entre sí.

Las pasiones son, según las consideraciones previas, una modificación de la mente marcada, en algunos casos, por una cierta relación con el placer o el dolor y, en otros casos, por esa misma relación y otros elementos adicionales, esto es, otras ideas. El primer caso da lugar a las pasiones directas y el segundo, por la mediación con esas otras ideas, hace surgir a las pasiones indirectas. Además, aun cuando todas las pasiones se relacionan con el placer o el dolor, algunas pasiones se sienten vivamente, son “violentas” según Hume, y otras son casi imperceptibles, por ello son designadas como “calmadas” (recordemos cómo esta condición fenomenológica surge de la teoría de las ideas de Hume, cf. supra p. 16). La siguiente tabla presenta esta primera distinción (cf. T 2.1.1 SB 276):

²³ Fate Norton explica que para Hume las pasiones tienen tanto dimensión temporal como dirección. Desarrolla el modelo de explicación de Hume que consiste en presentar cómo dos pares de pasiones se corresponden en sus impulsos y direcciones. Por ejemplo, el amor y el odio parecen ejercer un impulso similar sobre un agente hacia otro agente, pero en dirección opuesta: el amor de atracción y el odio de repulsión. De acuerdo con esto, «muchas pasiones están unidas [*conjoined*] con apetitos o deseos, y las pasiones así unidas con los deseos aparentarán tener mayor fuerza e incluso tendencias inesperadas», v. Fate Norton 2000, I64).

Pasiones directas	Pasiones calmadas	Pasiones violentas	Pasiones indirectas
	<p>Surgen directamente del bien o mal, del dolor o del placer.</p> <p>Son el sentimiento de belleza o fealdad hacia una acción, una composición [artística] y hacia los objetos externos (T 2.1.1.3 SB 276)</p>	<p>Surgen directamente del bien o mal, del dolor o del placer, <i>y de otras cualidades.</i></p> <p>Además de las anteriores, incluyen, ambición, vanidad, amor, esperanza, generosidad (2.1.1.4 SB 276-7).</p>	
	Incluyen deseo, aversión, pena, miedo, desesperación y seguridad (T 2.1.1.4 SB 276).	Ejemplos: Amor, odio, pena, alegría, orgullo y humildad.	

El sentimiento de belleza, y su contraparte, el de fealdad, surgen directamente de los apetitos básicos de placer y dolor. Hume parece decir, según este contexto, que las pasiones directas de pena, miedo, desesperación y seguridad, así como el deseo y la aversión, están trenzados con tales sentimientos. Esto parece conformar principios activos de acción, dirigidos por aquello que nos parece hermoso o feo, respectivamente. Los agentes queremos lo bello y rechazamos lo feo. En estos casos, la determinación natural hacia lo bello y lo feo produce voliciones y acciones marcadas fenomenológicamente por el deseo y la aversión. Bajo una descripción semejante podemos presentar a las otras pasiones directas.

Los sentimientos de belleza y fealdad modulan dichas pasiones de manera que las experimentamos con sus peculiares atributos. Nos dan miedo aquellas percepciones dolorosas de horror, como la posibilidad de sufrir un accidente; o nos agobia la desesperación cuando percibimos un dolor continuo del que no sabemos cómo liberarnos, como cuando nos enfrentamos a una enfermedad larga, o a las miserias de la pobreza, etc. Aun cuando estas descripciones puedan ser correctas, no parece que sean suficientes para explicar tales estados mentales. Como he dicho, se requiere de una cierta percepción, la de aquella fuente de horror, de miedo o de desesperación. En otras palabras, parece que aun en el caso de las pasiones directas la presencia de una idea es necesaria como objeto de la pasión. Experimentamos estas pasiones ante una idea de enfermedad, de dolor, etc., además, también es necesario decir que el deseo y la aversión sean *acerca* de algo. Si esto es correcto, entonces, como reconoce Hume, «la distinción entre pasiones está lejos de ser exacta» no solo en el sentido en que lo dice él, entre pasiones calmadas y violentas, sino también

entre pasiones directas e indirectas. Hume podría responder a esta cuestión recordándonos que toda percepción está marcada por el placer o el dolor, de modo que podría decirse que su punto sobre las pasiones directas es que percibimos a ciertas ideas de manera placentera o dolorosa pero también bajo la impresión de belleza, repulsión, esperanza, miedo, etc. En el caso de los sentimientos de belleza o fealdad parece que la explicación tendría que decir que la percepción de aquello, impresión o idea, que experimentamos como placer o dolor es por sí misma la belleza o la fealdad. Esta precisión refuerza el pensamiento de que incluso para Hume las pasiones tienen contenido representacional, esto es, son acerca de algo. Parece, pues, que tenemos que reconocer que las pasiones, en efecto, surgen directamente del placer o el dolor, pero éstos son también una cierta percepción, la de una impresión o de una idea. Las pasiones, de este modo, son percepciones o representaciones placenteras o dolorosas de ciertos objetos.

Con base en esta observación Hume explica que nuestra búsqueda de la belleza, en las acciones, en el arte o en los objetos externos se origina directamente, aunque muchas veces de manera imperceptible, en el placer que la belleza nos produce. Correspondientemente, la fealdad nos resulta repulsiva y motiva que nos alejemos de aquellos donde la percibimos. En otras palabras, la búsqueda de la belleza está determinada por un cierto placer, distintivo de la naturaleza humana, y no por la razón o algún otro interés motivado cognitivamente. Esta es una determinación muy importante que habrá de mostrar más adelante su relevancia para explicar la motivación de las acciones morales (cf. T 3.3.1.6-10 SB 575-578; EPM V.1 81-82; cf. Wiggins 1993, 334-336, Wiggins 2006, 78-79 y Brand 1992, 24). Como veremos, esta determinación opera unida a la simpatía, por lo que tenemos que reunir más elementos de la teoría de las pasiones para explicar adecuadamente este deseo de belleza (v. abajo §6.2). Con esto en mente, consideraré enseguida la distinción entre pasiones directas e indirectas.

2.2. Las ideas, las sensaciones y las determinaciones de las pasiones directas e indirectas

De acuerdo con McIntyre, la aportación más destacada de Hume a la teoría de las pasiones es, precisamente, la distinción entre pasiones directas e indirectas.²⁴ La importancia de esta distinción se aprecia si adoptamos una perspectiva amplia sobre la relación del placer y el dolor con las

²⁴ Según McIntyre, identificar a las pasiones indirectas constituye una de las principales aportaciones de Hume, especialmente por contraste con la concepción racionalista de las pasiones (McIntyre, p. 209). Un párrafo más abajo digo algo para marcar la diferencia con la concepción racionalista de Descartes sobre las pasiones.

pasiones. En otro contexto, explicando cómo surgen las virtudes y los vicios naturales, Hume detalla las operaciones causales determinadas por el placer y el dolor:

El impulso fundamental o principio motor de la mente humana es el placer o el dolor; cuando estas sensaciones desaparecen de nuestro pensamiento y sensibilidad, somos casi totalmente incapaces de experimentar pasión o acción, deseo o volición. Los efectos inmediatos del placer y el dolor son los movimientos mentales de acercamiento y evitación, que se presentan como volición, deseo y aversión, tristeza y alegría, esperanza y miedo, según que el placer o el dolor cambien de situación y resulten probables o improbables, ciertos o inciertos, o sean considerados como estando fuera de nuestro poder por el momento (T 3.3.1.2 SB 574).

Este pasaje es prácticamente un resumen de la teoría de la motivación de Hume (cf. Fate Norton 2000, 192). De acuerdo con él, un motivo para actuar tiene su impulso original en el placer o el dolor. De manera inmediata estos apetitos dan lugar a diversos movimientos mentales, como la volición, los sentimientos calmados de belleza y fealdad, los deseos y las pasiones.²⁵ Enseguida ciertos juicios intervienen y modulan este proceso: "...según que el placer o el dolor cambien de situación y resulten probables o improbables, ciertos o inciertos, o sean considerados como estando fuera de nuestro poder por el momento". De manera que el proceso motivacional, que inicia con un apetito y es así no cognitivo, puede ser dirigido, impulsado o extinguido por un juicio (cf. T 2.3.3.4-6 SB 415-416).²⁶ En otras palabras, en su explicación sobre la motivación Hume reconoce que los juicios tienen un papel *instrumental* que consiste en servir a la satisfacción de los apetitos originales. Desarrollaré en el siguiente capítulo estas ideas sobre la motivación. Por ahora me detendré más en la relación causal entre esos apetitos y las pasiones indirectas.

En este caso, la relación del apetito de deseo o de aversión con *otras cualidades* explica al surgimiento de las pasiones indirectas. Tales cualidades, me parece, además de marcar la diferencia con las pasiones directas, determinan que los pocos elementos de la mente que estamos considerando causan unas pasiones en vez de otras. Hume centra su análisis de las pasiones indirectas en un grupo modélico de pasiones que, además, forman parte de la constitución moral de los agentes: el orgullo, la humildad, el amor y el odio (cf. Alanen, 186).²⁷ Su análisis se desarrolla considerando a las pasiones por pares, tal como lo hizo antes Descartes en su tratado

²⁵ Examinando las ideas de Hume sobre la motivación, en el siguiente capítulo, consideraré las implicaciones de esta idea, a saber, la de que la relación directa entre un deseo y la voluntad. Hume parece, en efecto, definir a la voluntad como la percepción directa de un deseo (cf. T 2.3.1.2 SB 399).

²⁶ La referencia es al argumento central de la teoría de la motivación y de la moralidad de Hume. El papel de la razón, modulando al impulso activo de las pasiones, coincide con esta observación de Hume sobre el proceso causal que inicia con los apetitos. Dedicaré, en el siguiente capítulo, mucha atención a este argumento.

²⁷ Respecto del orgullo y la humildad dice Alanen que éstas juegan un papel central en la formación de la imagen propia y la identidad personal, y son también cruciales para la moralidad.

sobre las pasiones. De este modo, resaltan las afinidades y diferencias del modelo de Hume con el influyente modelo mecánico-hidráulico racionalista de Descartes (cf. Alanen, 180-181). Entre las afinidades me parece que una importante es la observación de que las pasiones surgen de estímulos semejantes, pero que de acuerdo con el «bien» o «mal» que percibimos en ellos su *dirección* es opuesta. Esto determina que el amor y el odio, el orgullo o la humildad, sean opuestos en un espectro de estímulos semejantes. Así, Descartes dice a propósito del amor y el odio:

Todas las pasiones que preceden, pueden ser excitadas en nosotros sin que de ningún modo percibamos si es bueno o malo el objeto que las causa. Pero cuando se nos presenta una cosa como buena con relación a nosotros, es decir, como conveniente, esto hace que sintamos amor hacia ella; y si se nos presenta como mala o dañosa, esto nos excita al odio (Descartes, *Las pasiones del alma*, Art. 56).

La idea de que todas las pasiones surgen de la percepción del bien o del mal es el principio explicativo más importante, me parece, al que recurre Descartes en su estudio sobre las pasiones (v., entre otros, Arts. 52, 57, 61, 64, 65, etc., del libro citado). Descartes considera que los objetos no son por sí mismos malos o buenos, sino que estas cualidades están presentes en la percepción de ellos. Esta percepción, en combinación con otros factores causales, como el deseo o la aversión, hacen surgir a las distintas pasiones (cf. Art. 57). La diferencia más notable entre su modelo mecánico-hidráulico y el de Hume reside en que Descartes cree que hay seis pasiones básicas (la admiración, el amor, el odio, el deseo, la alegría y la tristeza) y que todas las demás son compuestos o especies de éstas (cf. Art. 69). Hume, por su parte, no afirma que haya algún estado mental primitivo básico constitutivo de la mente. Hume diría que los estados mentales se forman a partir de unos pocos elementos básicos, las percepciones o pensamientos, y ciertas determinaciones que operan sobre ellos.²⁸ Examinaré cómo es este proceso en el caso de las pasiones indirectas.

En el caso del orgullo y la humildad, que Hume parece considerar como modelos explicativos del resto de las pasiones indirectas, el rasgo que las caracteriza es que su objeto es la idea del propio yo o de uno mismo (*self*) (cf. T 2.1.2 SB 277).²⁹ Estas pasiones se *refieren* tanto a dicho yo y a un cierto objeto con el cual se relaciona: un agente dado experimenta placer o dolor por una percepción de la cual él mismo es objeto y, además, algo placentero o agradable con lo

²⁸ Además de Hume, varios otros autores han defendido alguna variación de este económico modelo de la mente y de las pasiones, destacadamente Platón, Aristóteles, los estoicos, Spinoza y Kant (cf. De Sousa 2013)

²⁹ Descartes también reconoce que el objeto del orgullo y de la humildad puede ser uno mismo, pero no considera que sea ni su único objeto ni que tal objeto sea solo un elemento constitutivo del orgullo (cf. Art. 54).

que se relaciona. Esta relación es experimentada como una forma de placer, en el caso del orgullo o como una forma de dolor, en el caso de la humildad. Hume dice que estas distintivas sensaciones de placer o dolor “constituyen la esencia misma” de tales pasiones (T 2.1.5.4 SB 286) y por ello, según dice Alanen, algunos emotivistas han tomado a Hume como reduciendo las pasiones a ser meras sensaciones o emociones, sin contenido alguno. Pero de acuerdo con ella, “esencia” aquí no se refiere a las condiciones necesarias y suficientes que caracterizarían al orgullo o a la humildad, sino que es «una forma de enfatizar que la impresión afectiva es un componente necesario del complejo de pensamientos o ideas constitutivos de la pasión» (cf. Alanen, 187). En efecto, este complejo de pensamientos, sensaciones e ideas que constituyen a las pasiones sí parece *englobado* por la sensación de placer en el caso del orgullo y de dolor en el caso de la humildad. Al mismo tiempo, la idea del yo y otras ideas con las que se relaciona son indispensables para explicar a estas pasiones y ellas no son meramente sensaciones. Con todo, el yo no puede ser *la causa* de estas pasiones, porque ellas son entre sí «directamente contrarias» y como el yo es el objeto de ambas, se sigue que éste no puede ser su causa. De manera que es necesario distinguir entre el objeto y la causa de la pasión (cf. T 2.1.2.3-4 SB 277-278).³⁰ Sin embargo, el propio yo, el objeto de estas pasiones, no parece una entidad tan sólida. Parece que la relación causal de las pasiones con el yo contribuye a la formación del yo. Los principios de la acción y la motivación humana, desde esta perspectiva, parecen determinados por dicha relación causal. El argumento de Hume al respecto puede reconstruirse en los siguientes términos.

Hume afirma, en este contexto, que el «yo» no es más una “sucesión de ideas e impresiones relacionadas de que tenemos memoria y conciencia íntima”, de manera que si el yo no es la causa de las pasiones, pero sí su objeto, como estamos considerando, parece que la explicación sobre cómo ciertas sensaciones placenteras y dolorosas causan al orgullo y a la humildad será también relevante para explicar la idea de unidad asociada a la sucesión de ideas e impresiones que identificamos como el yo. Hume observa que el yo ha sido experimentado como objeto del orgullo o de la humildad por toda la naturaleza humana, en todas las naciones y épocas, y las únicas variaciones que cabe reconocer al respecto se deben a las diferencias particulares de temperamento y complejión (T 2.1.3.4 SB 280-281; cf. Davidson 1976, 753).³¹ Así, el hecho de

³⁰ Davidson, quien con su interpretación cognitiva de las ideas de Hume sobre el orgullo contribuyó mucho a su revitalización en el siglo pasado, señala, a propósito de este pasaje, que el principal problema para la teoría de las pasiones de Hume es explicar cómo éstas producen al yo (cf. Davidson 1976, 46).

³¹ Davidson, en el pasaje referido, indica también que el orgullo es universal.

que podamos reconocer al orgullo o la humildad revela ciertas cualidades de las operaciones mentales, que producen tanto al yo como a las pasiones indirectas, a saber: la «constancia y regularidad» (cf. T 2.1.3.2 SB 280). Éstas son cualidades naturales y primarias de la mente y, observa Hume, sin ellas no serían posibles ninguna otra cualidad secundaria, porque sin ellas “no habría fundamento alguno para la acción, ni podría ésta comenzar a ejercerse”. La constancia y la regularidad resultan entonces las cualidades constitutivas más básicas e irreductibles del «alma». Así, encontramos de nuevo cómo para Hume las determinaciones naturales de la mente causan a los distintos estados mentales y, en este caso, también a la idea del yo. Como vimos en el caso de las creencias sobre hechos, la regularidad es también una determinación que está detrás del principio de asociación y, por ello, determina a tales creencias (cf. supra p. 19). En este contexto, además del yo, dicha determinación explica al orgullo y a la humildad: “Y de este tipo es justamente la cualidad [original e irreductible] que determina el objeto del orgullo y la humildad” (cf. T 2.1.3.3 SB 280). El argumento de Hume, simplificando, reconoce que la constancia y universalidad de las manifestaciones del orgullo y de la humildad revelan los rasgos más básicos de las determinaciones mentales. Al parecer, algunos autores han señalado que este argumento es falaz y que comete petición de principio, sin embargo, a mí me parece que el argumento es *trascendental*, esto es, es un argumento que busca la mejor explicación para ciertos hechos observados.³² En este caso, los hechos observados son la universalidad de las manifestaciones de las pasiones de orgullo y de humildad, así como la idea de un yo, los cuales son explicados por Hume con base en la constancia y la regularidad. La manera en que estas determinaciones operan para producir estas pasiones requiere que descompongamos sus elementos más básicos.

La relación causal básica o «esencial» que explica al orgullo y a la humildad es entre su objeto, el yo, y una sensación placentera o dolorosa. Sin embargo, tales elementos suelen manifestarse con *independencia* de cualquier pasión, por lo que podemos afirmar que la relación entre ellos no es esencial o constitutiva de unos u otros. La tendencia a experimentar placer o dolor es propia de nuestra naturaleza y es una característica distintiva de las *impresiones* de sensación. Por su parte, la *idea* del yo parece un requisito de innumerables operaciones mentales. Sumando estos elementos encontramos que el orgullo o la humildad están constituidas por su

³² Cf. nota 11 de Félix Duque a su traducción al español del *Tratado* (en el contexto de T 2.1.3.3). De acuerdo con él, Glathe, en su *Hume's Theory of the Passions* (Berkeley, 1950), hace esta acusación de que Hume comete petición de principio.

objeto, el yo, la sensación de placer o dolor experimentada por el yo y dos propiedades que Hume *supone* como causalmente determinantes: la relación con el yo del placer o del dolor y la tendencia a producir placer o dolor «con independencia de la pasión». La pasión, entonces, se explica como una doble relación entre tales sensaciones placenteras o dolorosas asociadas y ciertas ideas relativas al yo. La pasión se deriva de tal doble relación mediante un proceso determinado por la *semejanza*, una de las determinaciones básicas de la mente según Hume: “Una de las ideas se convierte fácilmente en la idea correlativa, y una de las impresiones en aquella que se le asemeja y corresponde...” (T 2.1.5.5 SB 286-287; Davidson 1976, 747).³³ Por ejemplo, la idea de una casa bonita por sí misma me produce una sensación placentera, *independientemente* de cualquier pasión, pero si tal casa me pertenece, entonces la idea placentera de tal casa es *correlativa* con la idea de mí mismo, es decir, con mi yo y, al mismo tiempo, me produce una impresión de placer, la cual, a su vez, por *semejanza*, se convierte en la distintiva sensación placentera de orgullo. Necesitamos decir más acerca de por qué Hume parece pensar que solo la semejanza tiene las propiedades causales que pueden explicar al orgullo.

Con su modelo de la mente Hume propuso un novedoso sistema para explicar a los diversos estados mentales, con base en ciertas propiedades básicas de la naturaleza humana, a las cuales recurre en este contexto para explicar a las pasiones (cf. Alanen, 188). De acuerdo con tal sistema de asociación, «todas las impresiones semejantes están conectadas y tan pronto como surge una el resto la siguen inmediatamente» (cf. T 2.1.4.3 SB 283), por ello tanto la idea de placer (o dolor) de algo que poseemos correlativa con nuestro yo, así como la impresión placentera (o dolorosa) causan ideas y sensaciones semejantes. Hume no recurre a la contigüidad porque, como hemos visto, esta relación estrictamente carece de propiedades causales (cf. p. 19). Tampoco parece que experimentos orgullo meramente por asociación. La idea de las varias cosas que poseemos o con las que nos relacionamos no es suficiente para despertar esta pasión. Se requiere que esta relación involucre al yo de una manera específica y placentera. Si mi casa bonita es exactamente igual a la de mis doscientos vecinos difícilmente su posesión me hará sentir orgulloso. Pero si ninguno de mis compañeros de trabajo posee una casa como ésta, entonces, en su presencia, experimentaré orgullo. De esta observación se desprende que la idea de poseer una

³³ A Davidson la idea de que la semejanza produce al orgullo le parece una de las afirmaciones “más interesantes” de Hume. En efecto, de acuerdo con su paráfrasis, Hume sostiene que la propiedad por sí misma produce una impresión que “se asemeja y corresponde” al orgullo. Explicaré en seguida cómo podríamos entender esta relación de semejanza.

casa bonita se correlaciona de maneras cognitivas específicas con el objeto del orgullo, es decir, con el yo. Parece correcto suponer que esta correlación reúne elementos cognitivos que marcan que la relación sea de orgullo o no (cf. Davidson 1976, 746/348).³⁴ Con tales elementos me refiero a la percepción de aquellos rasgos de la relación con algún objeto que poseo que sí causan efectivamente al orgullo y que, *formalmente*, parecen caracterizar al orgullo. La semejanza, en consecuencia, operaría percibiendo tales rasgos formales del orgullo y activaría el placer distintivo del orgullo al mismo tiempo que parece que le da al yo una distintiva *personalidad*, esto es, aquellos motivos del orgullo que nos parecen exclusivos de un cierto agente y que nadie más comparte o muy pocos comparten. Para lograr esto, la semejanza simplemente requiere que el sujeto perciba en una idea o en una impresión de sensación aquellos rasgos formales peculiares para que inmediatamente los mecanismos de asociación relativos a la memoria le hagan recuperar ideas o sensaciones semejantes (cf. Baier 1991, 41). En efecto, no todos nos enorgullecemos por los mismos motivos y aquello que sí nos enorgullece hace resaltar, por encima de otros rasgos, nuestra peculiar *forma de ser*. Hume identifica estos rasgos como “limitaciones” de su sistema y su análisis se puede poner en los siguientes términos (cf. Brand 1992, 62).³⁵ El orgullo –o la humildad– es una pasión cognitiva que surge de la doble relación presentada arriba y de la superación de las siguientes limitaciones: 1. Cualquier cosa relacionada con uno mismo, que produzca placer o dolor, produce respectivamente orgullo o humildad; 2. El objeto que nos agrada o desagrade debe ser peculiar a nosotros o por lo menos compartido sólo por unas cuantas personas (p.e., dice Hume, la salud recuperada después de una larga enfermedad nos produce mucho placer, pero como la salud es compartida por muchas personas este placer no suele dar lugar al orgullo); 3. El objeto placentero o doloroso tiene que ser fácilmente discernible y obvio, no sólo para nosotros mismos sino para otras personas; 4. El tiempo en que estamos en posesión del objeto que nos da placer o dolor debe ser de una extensión considerable. Los objetos placenteros o dolorosos que poseemos nos producen orgullo o humildad durante prácticamente todo el tiempo que los poseemos y cuando estas pasiones cesan, por alguna razón, entonces tal vez sólo nos producen

³⁴ Davidson sostiene que la explicación del orgullo debe satisfacer ciertas condiciones formales, entre ellas, que esté involucrada una cierta *creencia* en relación con el objeto que causa nuestro orgullo. Así, no nos orgulleceríamos de poseer algo muy común, sino, como ha dicho Hume, algo que no es común y que valoramos positivamente.

³⁵ De acuerdo con Brand, los elementos que discuto a continuación muestran que las determinaciones naturales no son suficientes para explicar al orgullo. Creo que está forma de ponerlo es correcta y, como veremos en seguida, estas limitaciones permiten introducir elementos cognitivos que parecen influir sobre las ideas correlativas de aquello que se posee y el yo.

alegría o pena. 5. Las *reglas generales* tienen gran influencia sobre la pasión y la humildad (cf. T 2.1.6.3-8 SBN 291-293).³⁶ La semejanza, en resumen, opera reconociendo estos rasgos en las impresiones de los objetos con los que reconocemos que nuestro yo tiene una relación de posesión y de un modo «constante y regular» produce al orgullo. En un sentido de esta idea, parece entonces que la semejanza actualiza siempre la misma emoción de orgullo y, como veremos más adelante, éste parece realmente un rasgo distintivo de las pasiones, de acuerdo con el cual éstas tendrían un objeto *profundo* que aflora con cada manifestación de la pasión. Pero también hemos reconocido los rasgos formales de la pasión, de manera que parece también que podemos aceptar que hay un objeto *formal* de la misma. Por último, podemos reconocer también un objeto *inmediato*, esto es, aquel objeto percibido de manera inmediata y que parece desencadenar el complejo mecanismo de la pasión. El objeto profundo subyace o se esconde detrás de cada objeto inmediato. La percepción del objeto inmediato suscita que se reaviven una especie de memoria emocional con este tipo de objeto, con el cual el agente se ha relacionado a lo largo de su vida. El objeto formal recoge las características distintivas o definitorias de la pasión, por ejemplo, los rasgos que hacen al orgullo ser lo que es, independientemente de los objetos inmediatos o profundos que lo suscitan en un agente dado. Consideraré más adelante, de acuerdo con las ideas que Baier y Alanen al respecto, este complejo mecanismo de las pasiones y los objetos que parecen involucrar. Antes de esto quiero agregar más elementos cognitivos relativos a la idea del yo y determinantes de las pasiones.

Estos elementos pueden ser expresados reflexivamente por los agentes como creencias y juicios. Hume no dice que se trate de juicios y creencias en su pormenorización del proceso cuasi mecánico-hidráulico que según él da lugar y extingue a las pasiones. Sin embargo, necesitamos de tales elementos cognitivos para explicar las variaciones que determinan a dicho proceso. Los agentes podrían invocar a los elementos cognitivos, reflexivamente, o sencillamente contemplarlos, pasivamente. No me parece que una actitud sea menos común que la otra. Ahora bien, aunque Hume *no* plantea su explicación en términos cognitivos, su *diseño* de ocho

³⁶ Para Hume, las reglas generales son de dos tipos, las que obedecen al hábito y las que obedecen a la reflexión. Ambas tienen efectos similares sobre las pasiones y podrían determinarlas. Las reglas del hábito son aceptadas y puestas en práctica de manera pasiva por los agentes, pero las reglas de reflexión requieren que el agente tome alguna distancia de sus circunstancias y delibere sobre qué hacer (cf. T 1.3.10.11-12 SB Ap. 631-632; 1.3.13.11 SB 149; 3.3.1.20 SB 585). Este recurso, que en general me parece poco aprovechado por Hume, permite explicar ciertos procesos deliberativos que afectan a las pasiones. Bajo este proceso deliberativo, las pasiones son afectadas reflexivamente, precisamente por medio de las creencias que las determinan.

«experimentos» nos deja *observar* cómo la doble relación característica de las pasiones indirectas está determinada por tales elementos relativos al yo y su relación con ciertos objetos. En la siguiente tabla he esquematizado estos experimentos.

La primera columna se refiere a las ideas relevantes, los elementos cognitivos que me interesa resaltar, para explicar a la pasión. La segunda a las sensaciones de placer o dolor que surgen por la percepción de tales ideas. La tercera indica qué pasión surge de la doble relación esquematizada en las primeras dos columnas y que, me parece, determina a todo el proceso. Es fácil ver que el factor determinante decisivo consiste en una variación de los elementos de la primera columna y que el contenido de cada experimento puede describirse como una idea percibida o una creencia sostenida por el agente:

Experimento Primer término de la doble relación. En todos los casos, el objeto de la pasión es la idea de uno mismo.	Sensación Segundo término de la doble relación: placer o desplacer	Pasión
1. Un agente, en compañía de otros, se topa con un objeto común (p.e., “una vulgar piedra”) con el que no tiene ninguna relación.	Ninguna en particular: ni emoción, ni daño, ni ningún placer.	<i>Ninguna</i> (T 2.2.2.5 SB 333)
2. El mismo agente tiene ahora una relación con el objeto: la vulgar piedra le pertenece.	Ninguna en particular: la sola relación da lugar a una sensación de placer o desagrado. “Ningún objeto trivial y corriente, que sea incapaz de producir un placer o dolor separado de la pasión, podrá originar nunca las afecciones de orgullo o humildad, de amor y odio...” (T 2.2.2.6 SB 334)	<i>Ninguna</i> (T 2.2.2.6 SB 334)
3. La misma agente, en compañía de otros, se topa con un objeto agradable o desagradable, pero ni ella ni sus acompañantes tienen relación alguna con el objeto.	El agente o sus acompañantes posiblemente experimenten placer o desagrado, en todo caso, no parece que el objeto les pase «desapercibido».	<i>Ninguna</i> . Hay la sensación de agrado, pero falta la relación con uno mismo. En efecto, “un objeto que tenga solamente una de estas relaciones no podrá originar nunca una relación constante y estable [como una pasión]...” (T 2.2.2.7 SB 335).

Experimento	Sensación	Pasión
Primer término de la doble relación. En todos los casos, el objeto de la pasión es la idea de uno mismo.	Segundo término de la doble relación: placer o desplacer	
4.1 El mismo agente tiene ahora una relación con un objeto, p.e., la virtud (T 2.2.2.9 SB 336). 4.2 Uno de los acompañantes del primer agente tiene la relación con el objeto, la virtud (T 2.2.2.9 SB 336).	4.1 Placer 4.2 Placer	4.1. <i>Orgullo</i> (T 2.2.2.9 SB 336) 4.2 <i>Amor</i> . «Las afecciones se transforman y, dejando de ser orgullo, pasan a ser amor, pero también hay una doble relación, entre la idea del objeto y la impresión» (T 2.2.2.9 SB 336).

El aspecto causal mecánico de las pasiones indirectas *explica* cómo los mismos materiales que dan lugar a una cierta pasión dan lugar a cualquier otra, siempre y cuando los elementos involucrados varíen de *maneras específicas*, por medio de la semejanza, como hemos visto. Pero para que la variación tenga lugar me parece que es necesario que el agente tenga alguna *percepción* de cómo se plantean los elementos involucrados. Creo que esta percepción es suficiente para defender que el carácter cognitivo de las pasiones es el factor determinante principal en ellas. Podemos apreciar, nuevamente, el efecto de esta percepción en los experimentos restantes:

Experimento	Sensación	Pasión
Primer término de la doble relación. En todos los casos, el objeto de la pasión es la idea de uno mismo.	Segundo término de la doble relación: placer o desplacer	(Producida por transición de una cierta pasión a otra)
5. El agente anterior y su acompañante tienen relaciones por “lazos de sangre o amistad” (es hijo, hermano o tienen un «conocimiento mutuo largo y familiar»). Este compañero tiene una virtud o un vicio.	Suponemos que ante la vista de este compañero hay una impresión de amor u odio (T2.2.2.12 SB 338), es decir, la idea de este compañero se experimenta como una sensación de amor u odio –ya pasiones).	<i>Orgullo</i> si el compañero posee la virtud. <i>Humildad</i> si el compañero posee el vicio. La pasión de amor u odio, por “transición originada por las mismas causas da lugar al orgullo o la humildad” (T 2.2.2.13 SB 338). Las «mismas causas» son la idea del yo que ama u odia al compañero, más la relación de éste con un objeto, la virtud o el vicio. La transición es de la pasión original de amor u odio

<p style="text-align: center;">Experimento</p> <p style="text-align: center;">Primer término de la doble relación. En todos los casos, el objeto de la pasión es la idea de uno mismo.</p>	<p style="text-align: center;">Sensación</p> <p style="text-align: center;">Segundo término de la doble relación: placer o desplacer</p>	<p style="text-align: center;">Pasión</p> <p style="text-align: center;">(Producida por transición de una cierta pasión a otra)</p>
		<p>hacia la nueva pasión de orgullo o humildad.</p>
<p>6. El agente anterior y su acompañante tienen relaciones por “lazos de sangre o amistad” (es hijo, hermano o tienen un «conocimiento mutuo largo y familiar»). En este experimento, el compañero es objeto de amor u odio, pero no posee un objeto como la virtud o un vicio.</p>	<p>Tal vez amor u odio hacia el compañero (en buena medida porque hay lazos de sangre involucrados) (cf. T 2.2.2.14 SB 338), pero, al parecer, estos son los únicos sentimientos que se despiertan en el sujeto.</p>	<p><i>Ninguna.</i> “...este cambio en la situación [por relación con el experimento 5] produce una ruptura en la cadena entera y... la mente no pasa ya de una pasión a otra...” (cf. T 2.2.2.14 SB 338).</p> <p><i>Explicación:</i></p> <p>“...la imaginación y las pasiones, se ayudan mutuamente en sus operaciones cuando sus inclinación es similar, y también cuando actúan sobre el mismo objeto [lo cual no es el caso en este experimento]. La mente muestra siempre tendencia a pasar de una pasión a otra relacionada con ella, y esta tendencia se ve ayudada cuando el objeto de una de las pasiones está relacionado con el objeto de la otra. Los dos impulsos coinciden, haciendo que la transición sea suave y fácil. Pero si se diera el caso de que, a pesar de que la relación de ideas, estrictamente hablando, siguiera siendo la misma, <i>su eficiencia para hacer que la imaginación realizara la transición no pudiera ya ejercerse, es evidente que también su influencia sobre las pasiones debería cesar, en cuanto que ello depende por completo de esa transición</i>” (T2.2.2.16 SB 339-340, énfasis añadido).³⁷</p>
<p>7. El agente anterior y su acompañante tienen relaciones por “lazos de sangre o amistad”</p>	<p><i>Amor u odio</i>, en ambos sentidos, del agente al compañero sanguíneo y viceversa.</p>	<p><i>Amor</i> <i>Odio</i></p>

³⁷ Esta larga cita me pareció necesaria porque Hume indica explícitamente que la transición de una pasión a otra depende de que la imaginación, una facultad cognitiva, pueda relacionar adecuadamente los dos términos que distinguen a las pasiones indirectas, es decir, la idea del yo con la idea de otro objeto.

<p style="text-align: center;">Experimento</p> <p style="text-align: center;">Primer término de la doble relación. En todos los casos, el objeto de la pasión es la idea de uno mismo.</p>	<p style="text-align: center;">Sensación</p> <p style="text-align: center;">Segundo término de la doble relación: placer o desplacer</p>	<p style="text-align: center;">Pasión</p> <p style="text-align: center;">(Producida por transición de una cierta pasión a otra)</p>
<p>(es hijo, hermano o tienen un «conocimiento mutuo largo y familiar»). Tanto el agente –la idea de un yo- como el compañero son objeto de amor u odio y <i>ambos</i> poseen alguna virtud o algún vicio (cf. T 2.2.2.18 SB 341).</p>		<p><i>Orgullo</i></p> <p><i>Humildad</i></p> <p>“...cuando sentimos amor u odio por alguna persona, raramente permanecen las pasiones dentro de sus primitivos límites...” (cf. 2.2.2.18 SB T 341).</p> <p>La transición del amor u odio a las pasiones de orgullo o humildad es muy fácil y se refuerza además por la cercana relación familiar.</p>
<p>8. El agente anterior y su acompañante <i>no</i> tienen relaciones por “lazos de sangre o amistad, incluso, podría ser una persona muy distante. (cf. T 2.2.2.27 SB 346). El agente reconoce en sí mismo la posesión de un objeto, una virtud o un vicio.</p>	<p>Orgullo</p> <p>Humildad</p> <p>“...pues el objeto original es el yo” (T 2.2.2.27 SB 346).</p>	<p><i>Amor</i></p> <p><i>Odio</i></p> <p>“...es [evidente] que esta pasión [orgullo o humildad] se convierte en amor y [¿u?] odio, cuyo objeto es otra persona”.</p> <p><i>Explicación:</i></p> <p>Intuitivamente, no parece que experimentar orgullo por uno mismo pueda generar amor u odio por otra persona, precisamente porque “...<i>la imaginación pasa con dificultad de lo contiguo a lo remoto</i>” (T 2.2.2.27 SB 346). Además, en este caso la transición no es entre uno mismo y la otra persona. Sin embargo, se da la transición hacia el amor u odio porque “esa misma persona es la causa real de nuestra primera pasión, y por consiguiente está íntimamente ligada a esta pasión” (T 2.2.2.27 SB 346).</p> <p>Hume parece defender esta transición reconociendo que el orgullo o la humildad, la primera pasión, se originan por relación con alguien. Nos sentimos</p>

Experimento Primer término de la doble relación. En todos los casos, el objeto de la pasión es la idea de uno mismo.	Sensación Segundo término de la doble relación: placer o desplacer	Pasión (Producida por transición de una cierta pasión a otra)
		orgullos por nuestro trato o actitud hacia alguien.

Una prueba adicional a favor de la preponderancia del factor cognitivo en los procesos que dan lugar a las pasiones es apreciar cómo la imaginación, una facultad cognitiva, opera sobre tales ideas y contribuye así a despertar específicamente a ciertas pasiones (v. experimentos 6 y 8). Con todo, ¿podemos aclarar aún más qué clase de contenido cognitivo es éste? Por ejemplo, algunos autores consideran que las ideas, en estados mentales como creer, imaginar, desear, etc., funcionan para Hume como actitudes proposicionales (cf. Merrill, 59-60; Bennett 2001, 221-224).³⁸ Otros autores sostienen incluso que para Hume las pasiones también están constituidas por contenidos proposicionales (p.e. Baillie, 91; Stroud, 232). Autores como Davidson han defendido, *inspirados* por Hume, una teoría cognitiva de las pasiones suponiendo efectivamente que cierto contenido proposicional es constitutivo de las pasiones (v. Davidson 1976). Siguiendo a A. Baier y L. Alanen yo voy a defender que parece más plausible, como explicación general de las pasiones, renunciar a la idea de que las pasiones están constituidas por contenido proposicional y, en vez de ello, examinar de qué manera las pasiones son cognitivas y cómo ofrecer una adecuada caracterización de su contenido representacional.³⁹ En esta tarea, también las ideas de Hume son especialmente importantes.

2.2.1 El contenido representacional de las pasiones indirectas

En esta sección quiero defender que si pretendemos ofrecer una lectura satisfactoria de la teoría de las pasiones de Hume, entonces «necesitamos una concepción de la intencionalidad más amplia que aquella intencionalidad construida en términos de actitudes proposicionales» (cf. Alanen

³⁸ En detrimento de esto Bennett interpreta a Hume como compatible con una posición funcionalista, en virtud de «la frecuencia con que enfatiza el fuerte vínculo entre la creencia y la conducta», (cf. Bennett 2001, 227).

³⁹ La teoría cognitiva del orgullo de Davidson, de inspiración humeana, enfrenta varias dificultades. Entre ellas destaca que su tratamiento del orgullo como una pasión cognitiva con contenido proposicional no parece representar bien todos los casos de orgullo. Por ejemplo, su teoría no puede expresar atribuciones comunes orgullo como: “A está orgullosa *de que* fue elegida presidenta” o “A está orgullosa por haber sido elegida presidenta”, que no son actitudes proposicionales sino, como nos muestra Baier, oraciones con predicado nominal (v. Baier 1978, 28). Davidson, así, parece reducir injustamente varias atribuciones de orgullo a una forma que ajusta bien a una concepción limitada a las actitudes proposicionales y deja fuera otros elementos que parecen, en otros casos, necesarios para una correcta descripción del orgullo (v. p.e. Hansberg, 137).

2005, 117-118). En consecuencia, la tarea es establecer en qué consiste esta concepción amplia de la intencionalidad. Comenzaré presentando la teoría general de Baier sobre las emociones, enraizada en la filosofía moderna, particularmente en la obra de Descartes y Hume, porque me parece que su teoría satisface especialmente bien el requisito presentado arriba.

Esta teoría se desarrolla a partir del estudio de dos hechos sobre las emociones ya observados por Descartes. El primer hecho es que la música, tanto por medio de variaciones tonales como de ritmo, influye poderosamente sobre las emociones de quien la escucha (cf. Baier 1990, 3).⁴⁰ El segundo hecho es que algunas de las emociones tienen objetos “profundos”, esto es que «detrás del objeto inmediato permanecen los fantasmas de otros objetos con el cual el tipo específico de una emoción se ha relacionado a lo largo de la historia de una persona» (cf. Baier 1990, 4). El primer hecho es desconcertante porque parece sugerir que la música induce pensamientos sin contenido y sin objeto (cf. Baier 1990, 12). Por contraste, Baier asume que un rasgo distintivo de las emociones es que tienen objeto o que son acerca de *algo* –por contraste con los estados de ánimo que parecen ser acerca de *todo*, p.e., la nostalgia es una emoción, cuyo objeto son ciertas alegrías perdidas, mientras que la depresión, al no ser acerca de ningún objeto en particular, es un estado de ánimo (cf. Baier 1990, 4). Tampoco voy a dedicar más atención a la posibilidad de que existan emociones sin objeto, es decir, que no refieran a nada y que aun así conserven el estatus de emociones.⁴¹ Cabe recordar, a propósito, que el propio Hume, al afirmar

⁴⁰ Baier cita ampliamente el *Compendium Musicae* de Descartes, una obra que, de acuerdo con ella, ha sido tratada con negligencia (cf. Baier 1990, 3). Cabe advertir que Baier reconoce que los términos de pasión o emoción empleados por Descartes, Hobbes, Hume, Darwin o Freud (autores a cuyas obras remite en este ensayo) refieren más o menos a lo mismo y que son “inter-traducibles”. También en esto seguiré a Baier.

⁴¹ Excepto Hume, y sólo si se le interpreta como un exponente muy estricto de una teoría fisiológica de las emociones, no parece que ningún autor defienda seriamente esta posibilidad. Las emociones tienen objeto e incluso según para la teoría fisiológica James-Lange éste podría ser alguna modificación corporal, por ejemplo, percibir un aumento repentino en la frecuencia cardíaca (cf. La introducción de Calhoun y Solomon a su texto de 1989). El problema interesante parece provenir de las objeciones típicas a las teorías cognitivas de las emociones. Estas teorías sostienen que las emociones son (incluso que están constituidas) por un juicio evaluativo o por alguna proposición adecuada. Por ejemplo, el estado ‘tengo miedo de que un alacrán me pique’ está constituido por la emoción estricta –el miedo– y por un elemento cognitivo adecuado ‘un alacrán me puede picar’ (es adecuado porque el piquete de un alacrán suele ser dañino, si no lo fuera, nuestro miedo *no estaría justificado*). Así, la emoción del miedo *es* un tipo de actitud proposicional. Pero si aceptamos esto, tenemos que reconocer que ciertos conceptos y proposiciones están involucrados con las emociones. Ahora bien, las objeciones típicas reconocen que tanto los niños muy pequeños, como los bebés, y varios animales, experimentan también miedo. ¿Estamos dispuestos a atribuirles, como sugiere la pregunta de Baier, emociones sin contenido o sin objeto? Es decir, parece claro que los niños muy pequeños y los animales carecen de pensamiento proposicional y tal vez incluso carecen de conceptos. Aun así reconocemos que experimentan emociones. La estrategia de Baier para defender una interpretación cognitiva de las emociones y responder a estas objeciones típicas descansa, como veremos, en una comprensión peculiar del concepto de objeto de una emoción, es decir, del concepto de intención.

que las pasiones son existencias originales que carecen de cualquier cualidad representativa, *parece* sostener una posición de este tipo (cf. T 2.3.3.5 SBN 415). Sin embargo, precisamente contra esta posición pretendo defender una interpretación cognitivista de su teoría de las pasiones. En la siguiente sección examino cómo la ampliación de Alanen a la teoría de Baier permite ofrecer una interpretación de este pasaje del *Tratado* consistente con la teoría cognitiva que estoy desarrollando. Consideremos con más detalla el primer hecho mencionado.

Supongamos que escuchamos una pieza musical y las emociones que nos sobrevienen son del tipo de la tristeza, del asombro, de la nostalgia. Al mismo tiempo, ciertos pensamientos son despertados. Recordamos a una persona, un lugar, queremos ver a alguien, pensamos en cosas específicas. Estos pensamientos nos sobrevienen con la música. En el tren de pensamientos que cabe suponer que tendríamos *sin* la pieza musical que escuchamos, no viajaría ninguno de estos pensamientos. Éstos están relacionados con la música, aunque no vienen con ella, porque nuestras historias respectivas, diferentes de por sí, lo impiden y porque la música no parece representativa en el sentido en que mis pensamientos lo son. De modo que tienen que provenir de mí, de mi historia, y la música que «influye poderosamente sobre mis emociones» los ha provocado de alguna manera. Al mismo tiempo, si escuchara en su lugar otra pieza seguramente mis emociones serían completamente diferentes. Tal vez no serían emociones de tristeza, sino semejantes al regocijo, al placer, al entusiasmo. Nuevamente, tendría pensamientos específicos suscitados por la pieza, pero serían muy distintos, en ese momento, a los pensamientos *sin* la pieza o *por comparación* con la primera pieza musical. Entonces, tiene que haber una relación inteligible entre ciertas piezas musicales y las emociones suscitadas por ella. ¿Cómo explicamos esta relación?

Los pensamientos específicos y las emociones no se refieren a la pieza musical o, en este caso, la pieza musical no es el objeto de mi emoción. Tampoco mis pensamientos y emociones están *contenidos* por las piezas, *i.e.*, no es cómo si leyera en ellas las emociones que experimento. ¿Qué sucede entonces? Al parecer, la respuesta tiene que provenir de la relación que hay entre mi experiencia emotiva del momento y mi historia emocional. Baier sostiene que proviene «del hecho de que las emociones en cuestión tienen objetos formales y profundos» (cf. Baier 1990, 12). Un rasgo distintivo de las emociones, según observamos antes a propósito de la determinación de la semejanza, es que tienen objeto, pero este objeto, en algunas emociones, puede ser formal (cf. arriba p. 49). Baier observa, para defender esta caracterización, que es inconcebible tener una sola

emoción, pues tener una emoción es tener un rango de emociones (cf. Baier 1990, 5).⁴² Los estados emocionales son complejos, moleculares, que parecen presentarse en paquetes más o menos discernibles: en el enojo hay algo de orgullo, de indignación, de agitación, de deseo de retribución y una cierta idea de uno mismo. Estos elementos constituyen la representación del complejo emocional, pero no todos ellos se refieren al objeto inmediato que la causa. Si me enojo porque alguien me avienta en la fila, varios de estos elementos entran en juego, pero parece evidente, además, que el enojo recrea ciertas condiciones previas del enojo. Sabemos que la gente no debería aventarnos, que la desconsideración es una descortesía, que la reacción apropiada a un atropello es la indignación, etc. Sabemos de alguna manera todo esto, hemos experimentado anteriormente este conjunto de respuestas emotivas, el cual se activa inmediatamente con las percepciones (incluso creencias) y evaluaciones directamente involucradas en la emoción actual. De este modo, el enojo actual tiene un contenido representativo proveniente de experiencias pasadas, y no directamente relacionado con la experiencia presente que lo suscita. Esta forma de entender a la representación parece correcta porque indica cómo una cierta respuesta es adecuada para una situación dada. Así, si es adecuado enojarse en esta situación, entonces la representación involucrada es correcta (cf. Baier 1990, 6). La experiencia emocional pasada recreada en el presente señala que la representación es adecuada. Así, el objeto del enojo es también formal, debido a que es representacional, si bien no representa a ninguno de los objetos presentes ante el agente. En la teoría de Baier este rasgo proviene del hecho de que, según dicha teoría, el objeto de las emociones también puede ser un objeto profundo.

En la historia de cada agente, el objeto profundo se encuentra enterrado bajo capas de emociones. “Las emociones —dice Baier— son estados de la mente cargados de historia” (Baier 1990, 18). Si queremos encontrar el objeto profundo de una emoción particular, la teoría indica que tendríamos que buscar capa tras capa de la historia de un agente hasta determinar el objeto original. Sólo así podríamos explicar correctamente la intencionalidad de una emoción determinada (Baier 1990, 14-15). El carácter formal de un objeto se descompondría, con cada capa de la historia de la emoción, hasta llegar al objeto profundo mismo. Sin embargo, nos advierte Baier, «tenemos que ser cautos en hacer equivaler al objeto “profundo” de una emoción

⁴² Baier apoya esta idea con un pasaje del *Tratado* donde Hume imagina una criatura simple que tiene una sola percepción, sed o hambre. Al considerar esta posibilidad, pronto queda claro que este ser sólo consistiría en la percepción misma de sed o de hambre, y carecería por completo de la noción de *yo* o de *sustancia*. Hume pronto identifica al yo con el conjunto de percepciones de la mente. (cf. T *Apéndice* 16-17 SBN 634-635).

con el objeto experimentado en la primera infancia de este tipo de emoción» (cf. Baier 1990, 14). En efecto, más importante que esta posible arqueología de las emociones, en la teoría de Baier es:

...localizar a las emociones en relación con los deseos y las creencias, no sólo en sí mismas, sino también en su dimensión de criticabilidad, ofrecer un modo notable de ver si y cómo las emociones (o tipos particulares de emoción) pueden ser “falsas”, “irracionales”, “desorientadas” (o “desorientadoras”), cómo sus fallas se comparan con el tipo de excelencias y fallas a las que las creencias y las secuencias de creencias son propensas, y así sucesivamente (Baier 1990, 15).

De modo que junto con el objeto formal de la emoción, que a su vez presupone un objeto profundo de las mismas, Baier quiere defender que una emoción será plenamente evaluativa si los objetos mencionados coinciden con el objeto *aparente*, es decir, con el objeto que desencadena a la emoción actual en un momento dado (cf. Baier 1990, 18). En otras palabras, la relación de las emociones con las creencias y los deseos es evaluable en términos racionales (por su falsedad, irracionalidad o desorientación), y para evaluar de este modo a las emociones *sólo* necesitamos establecer correctamente la relación actual de la emoción con ciertas creencias y deseos con los que se relaciona *en tanto que* emoción aparente. La manifestación de una emoción, en un momento dado, tiene una intencionalidad compleja y presupone tres objetos distintos pero convergentes, el objeto formal, el objeto profundo y el objeto aparente. No necesitamos, en consecuencia y *generalmente*, realizar una arqueología de la emoción para desentrañar al objeto profundo porque las condiciones de racionalidad ante el objeto aparente son *suficientes* para entender a cabalidad la manifestación de emoción —sin menoscabo de los objetos presupuestos (cf. Baier 1990, 15).⁴³ Pero ocasionalmente sí necesitamos de hacer alguna exploración del objeto profundo para determinar fallas en nuestras respuestas racionales.

Supongamos que me avientan en la fila y yo reacciono a esto con enojo. Sin embargo, pronto se aprecia que fue un accidente y que, ninguna creencia *justifica* racionalmente que continúe enojado, pero yo *continúo* enojado y de una manera francamente excesiva. ¿No parece que, si tengo algún tipo de interés en el bienestar de quienes me rodean y en el mío propio, debería examinar atentamente, sólo o con algún tipo de ayuda, la *causa* real de mi enojo? En casos

⁴³ Baier parece señalar (p. 15) que una explicación adecuada de la intencionalidad de una emoción tiene que dar lugar al objeto formal, profundo y aparente de la misma y “debería preferentemente ser capaz de indicar si el objeto profundo es el mismo que la aparición más remota del mismo”. Pero agrega enseguida que no sólo necesitamos escenarios paradigmáticos (que podrían provenir de una arqueología profunda de la emoción, como la práctica el psicoanálisis), “sino también escenarios finales donde el tipo de una emoción realmente llega a ser propia, donde encuentra su realización plena”. Me parece, en efecto, que tal como continua argumentado Baier, la realización plena de una emoción consiste en la satisfacción de las condiciones de evaluación racional que provienen de la convergencia de los tres objetos de la emoción y una relación adecuada con creencias y deseos relativos al objeto aparente.

semejantes a éste parece que es necesaria una exploración atenta del objeto profundo de mi enojo, pues la emoción actual del enojo no es despertada por el objeto aparente del mismo. Es posible que el objeto profundo sea muy remoto y que posiblemente esta exploración me remonté hasta la infancia, pero no es necesario. Podría tener algún tipo de rencor relativamente reciente. Estoy enojado porque me estafaron comprando algo y ahora tengo oportunidad de desahogarme, etc. En la literatura sobre emociones, casos de este tipo son identificados como emociones recalcitrantes, esto es, cuando existe un cierto desajuste entre las creencias que acompañan a la emoción y la emoción misma, de tal manera que la emoción persiste a pesar de que las creencias no la justifican. Las creencias, como en el ejemplo que considero arriba, señalan que no hay razón para continuar enojado, a pesar de lo cual yo sigo *recalcitrantemente* enojado (cf. p.e., D'Arms y Jacobson 2003, 129). La irracionalidad en la respuesta ante el objeto aparente muestra la importancia del objeto profundo en la explicación de las emociones recalcitrantes. En muchos de estos casos de respuesta irracional ante el objeto aparente, parece que tendríamos que buscar al objeto profundo para poder *explicar* —no *justificar*— dicha respuesta (cf. Baier 1990, 17). Un aspecto de la teoría de las pasiones de Hume es útil para entender esta relación del objeto profundo con nuestras respuestas emocionales recalcitrantes.

Tal como nos recuerda Baier, un rasgo que caracteriza fuertemente a las emociones es su inercia, es decir, su tendencia a sobrevivir a las creencias que les ofrecen sustento cognitivo (cf. Baier 1990, 16). Baier cita ahora un pasaje clave de la teoría de las pasiones de Hume que, en vista de su relevancia actual y posterior para la interpretación de Alanen, cito extensamente:

Si examinamos ahora la mente humana, encontraremos que, por lo que respecta a las pasiones, su naturaleza no es la de un instrumento de musical de viento, en donde, una vez recorridas todas las notas, cesa inmediatamente el sonido en cuanto se deja de soplar, sino que se asemeja más bien a un instrumento de cuerda, en el que, después de cada golpe de arco, las vibraciones siguen reteniendo algún sonido, que va desvaneciéndose gradual e imperceptiblemente. La imaginación es en extremo rápida y ágil, pero las pasiones son lentas y obstinadas (T 2.3.9.12 SBN 440; cf. Baier 1990, 16).

Hume con esto señala que las pasiones continúan ejerciendo sobre la mente la influencia original con la que fueron despertadas. Cuando estoy enojado, esta emoción ejerce una fuerza poderosa sobre la mente, la cual, dada la inercia de la emoción, continúa manifestándose, lo que puede dar lugar al tipo de emoción recalcitrante que estamos considerando. Las creencias, por su parte, son descartadas sin mayor resistencia por parte del agente una vez que se acepta su falsedad o falta de relevancia. Es verdad, digamos, que me aventaron en la fila, pero también es verdad que fue por

accidente, no deliberadamente, de modo que la creencia verdadera ‘me aventaron en la fila’ ha perdido su relevancia para justificar mi enojo. En el caso de las emociones no recalcitrantes, esto es suficiente para que la emoción desaparezca, pero en el caso de las emociones recalcitrantes, el objeto profundo mantiene su influencia sobre la mente y sobre la percepción del objeto aparente. Con todo, aun en los casos de las emociones no recalcitrantes, la fuerza inercial de las emociones ejerce alguna influencia. La diferencia reside en la fuerza relativa de ambos tipos de emociones para lidiar con el objeto aparente de la emoción.

La creencia ‘me aventaron en la fila por accidente’ afecta a la creencia ‘me aventaron en la fila’ de modo que sustrae a ésta la relevancia que tenía para ser una razón para mi enojo. Del mismo modo, las emociones se relacionan entre sí y modifican su carácter (cf. Baier 1990, 17). Es posible que ante el reconocimiento de que fui aventado por accidente mi emoción pase de enojo a desconcierto o a otra forma de indignación (¡deberían arreglar este piso!). Así, «entender una emoción no es separable de entender cómo surgió, tal como entender una creencia se entiende a partir de cómo una persona llegó a tenerla» (cf. Baier 1990, 17). Bajo estas apreciaciones, gana fuerza la afirmación de Baier de que las emociones son estados cargados de historia (cf. Baier 1990, 18). Recapitulemos brevemente los elementos de la teoría de Baier sobre las emociones que nos permiten lidiar con la supuesta contradicción en Hume que me interesa discutir.

La teoría de Baier, que se apoya en alguna medida en la idea humeana de que las pasiones se distinguen por su inercia, es decir, por la capacidad que tiene el *objeto profundo* de las emociones de seguir ejerciendo su influencia sobre otros estados mentales, admite sobre esta base que las emociones tienen un tipo amplio de intencionalidad o de cualidad representativa. El objeto profundo de la emoción subyace en las capas de la historia emocional del agente y se manifiesta muy claramente en las emociones recalcitrantes, aunque su carácter latente o inercial opera constantemente sobre la mente. Además del objeto profundo de la emoción, las emociones tienen un *objeto formal* (que reúne los distintos elementos que se expresan en una emoción, como el propio objeto profundo, rasgos de carácter, una cierta idea del yo, etc.), el cual no se refiere directamente al *objeto aparente*, es decir, el objeto que, en última instancia, es la causa inmediata de la emoción. En resumen, la representación o intencionalidad de una emoción consiste en una noción compleja y amplia que está constituida, en cada caso, por el objeto profundo, el objeto formal y el objeto aparente de la emoción. Examinaré a continuación la interpretación de L. Alanen de la teoría de las emociones de Hume. Su importancia para mi argumento reside en que

mantiene los elementos esenciales de la teoría cognitiva de Baier y presenta de qué manera las pasiones tienen contenido intencional sin ser, necesariamente, representativas. De este modo logra explicar la idea de Hume de que las pasiones son existencias originales sin ninguna cualidad representativa.

2.2.2. La interpretación de Alanen de la teoría de las pasiones de Hume

De acuerdo con Alanen, una correcta interpretación de la teoría de las pasiones de Hume requiere de una concepción de intencionalidad más amplia que la construida en términos de actitudes proposicionales (cf. Alanen 2005, 118). Éste, como hemos visto, es un requisito que proviene de la teoría de las pasiones de Baier y que Alanen suscribe. Ahora bien, según Alanen, Hume nos ofrece una concepción de intencionalidad amplia. En primer lugar, Alanen explica a la representación en las pasiones distinguiendo entre la forma en que una idea refleja a una impresión y la forma en que una pasión refleja a una impresión. Las ideas sí representan a las impresiones, pero las pasiones, en tanto que existencias originales, no las representan, pero sí las reflejan. Como impresiones de reflexión, reflejan ideas. Al parecer, en esta distinción están presentes dos metáforas diferentes: una es *visual*, la relación de copia, e ilustra cómo las ideas reflejan a las impresiones. La otra es *musical* e ilustra cómo las impresiones secundarias son resonantes de las impresiones primarias y de sus ideas, prolongando y reforzando sus efectos como si por medio de vibraciones y ecos en la persona que percibe los acordes correctos o el punteo de una cuerda (cf. Alanen 2005, 118). Como veíamos arriba, Hume en efecto concibe que la naturaleza de las pasiones es más como la de un instrumento de cuerdas que continúa reverberando después de haberlo tocado que como un instrumento de viento que sólo suena mientras soplamos a través de él (cf. T 2.3.9.12 SBN 440). Ésta es la idea que Baier desarrolla para establecer que las pasiones tienen un objeto intencional profundo y, como vemos, Alanen desarrolla su interpretación con base en ella. Además, esta tesis de la reverberación tiene otro sentido, relacionado con la capacidad de reflejar las pasiones de otras personas.

La teoría del objeto profundo, entendida como la noción de que el objeto intencional de las pasiones no es solamente el objeto inmediato presente, sino un objeto remoto, ubicado en algún lugar de la historia de la pasión, tiene también una contraparte dirigida hacia otras personas y es la base de la teoría de la simpatía de Hume —y, por extensión, de la moralidad de Hume. Dos pasajes del *Tratado* son claves para respaldar esta idea. Consideremos el primero:

En general podemos decir que las mentes de los hombres son espejos unas de otras, y esto no sólo porque cada una de ellas refleja las emociones de las demás, sino también porque la irradiación de las pasiones, sentimientos, y opiniones puede ser en muchas ocasiones reverberada, e ir decayendo por grados imperceptibles (T 2.2.5.21 SBN 365).

De acuerdo con Alanen, en efecto, nuestras pasiones afectan a otras personas y sus respuestas emocionales respectivas, a su vez, nos afectan y suscitan nuevas respuestas emocionales (cf. Alanen 2005, 119). De este modo podemos explicar cómo las emocionales reflejan las pasiones de otras personas. Baier, al respecto, dice que todas las emociones son acerca de nuestra condición humana (cf. Baier 1990, 19). La metáfora del espejo explica también distorsiones en las respuestas emocionales recíprocas. Pienso, por ejemplo, en cómo experimentamos frustración como respuesta a un gesto de desdén de otra persona, o indignación o depresión. No siempre una respuesta emocional se corresponde fielmente con la emoción que refleja. Esta posibilidad de que las pasiones humanas se reflejen mutuamente muestra otro aspecto formal de la naturaleza de las emociones. Examinemos esto considerando el segundo pasaje clave mencionado.

Al examinar la naturaleza de la simpatía, Hume nos ofrece otros aspectos de la naturaleza reflejante de las pasiones:

Las mentes de los hombres son similares en sentimientos y operaciones y no hay ninguna que sea movida por una afección de la que, en algún grado, estén libres las demás. Del mismo modo que cuando se pulsan por igual las cuerdas de un instrumento el movimiento de una se comunica a las restantes, así pasan fácilmente de una persona a otra las afecciones, originando los movimientos correspondientes en toda criatura humana (T 3.3.1.7 SBN 575-576).

La metáfora de la reverberación de las pasiones se muestra ahora en relación con la simpatía. Las mentes humanas operan igual y experimentan los mismos sentimientos. La reverberación en una mente encuentra en otras mentes condiciones para ser recibida de acuerdo con su *ritmo* distintivo. La alegría, el miedo, el entusiasmo, la indignación, el enojo, entre otras, son emociones que se *contagian* fácilmente (cf. T 2.1.11.2 SBN 317; Stroud, 286). Esta observación parece requerir, para ser explicada, de la siguiente condición formal: se necesita de una cierta disposición natural para responder emotivamente a *cualquier* emoción de otra persona. De acuerdo con Hume, la simpatía es esta disposición natural (cf. T 2.1.11.2 SBN 316; 3.3.1.7 SBN 575-576). Pero no parece que la simpatía misma sea un sentimiento determinado (cf. Stroud, 287). En efecto, de acuerdo con el modelo causal de la mente de Hume, la simpatía tiene la función de producir pasiones al transformar a las ideas en impresiones:

Cuando se infunde por simpatía una cierta afección, al principio es reconocida solamente por sus efectos y signos externos, presentes en el gesto y la conversación, y que dan una idea de esa pasión. Esta idea se convierte entonces en una impresión, adquiriendo de este modo tal grado de fuerza y

vivacidad que llega a convertirse en pasión, produciendo así una emoción idéntica a la de una afección original (T 2.1.11.3 SBN 317).

Ésta es una descripción no metafórica de la reverberación que existe entre las pasiones. Hume presenta aquí cómo es que las pasiones son impresiones de reflexión, causadas por una relación con una idea, y cómo las pasiones no son copias de la idea, sino un reflejo de la idea que tiene propiedades de una impresión, es decir, en sus términos, que tiene la fuerza distintiva de las impresiones para influir sobre la mente. En la última sección de este capítulo examinaré el papel de la simpatía en la teoría de las pasiones de Hume.

2.2.3. Simpatía e imaginación en la teoría de las pasiones de Hume

La observación de que las pasiones se reflejan unas a otras *requiere* de una explicación que muestre cómo una cierta disposición de la mente da lugar a dicho reflejo. Esta disposición es la simpatía, pero necesitamos distinguirla de la asociación de semejanza, la determinación básica de la mente que parece realizar la misma operación. Cabe aclarar que no parece necesario distinguirla de la contigüidad, la otra asociación básica entre ideas porque su operación no relaciona ideas por semejanza, esto es, por alguna forma de reflejo. La contigüidad, como hemos visto, relaciona a las ideas que se presentan a la mente en intervalos de tiempo o espacio contiguos, de modo que la mente pasa de una a otra sólo por su cercanía, no por las características de las ideas (v. supra p. 19). De acuerdo con Hume, parece que la simpatía es una determinación adicional de las pasiones que opera sobre las determinaciones básicas de la semejanza y la contigüidad (cf. T 2.1.11.4-6 SB 317-318). Hume parece pensar en la simpatía como un factor adicional que determina el surgimiento de las pasiones no directamente por la percepción de una impresión y una idea placenteras asociadas con nuestro yo, sino por el reconocimiento de "...la opinión de los demás, y que tiene una influencia igual sobre las afecciones." Esta influencia se manifiesta en la consideración que tenemos por nuestra reputación, nuestro carácter o nuestro nombre (cf. T 2.1.11.1 SB 316; v. también: T 2.3.6.8 SBN 427 y 3.3.2.5 SBN 595). La opinión de los otros, o el punto de vista de otras personas es, desde el punto de vista de Hume, la percepción de una idea y el papel de la simpatía consiste en «convertirla en una impresión» de modo que la idea adquiera "tal grado de fuerza y vivacidad que llega a convertirse en una pasión, produciendo así una emoción idéntica a la de una afección original" (T 2.1.11.3 SB 317). Así, en este caso las pasiones no surgen por la observación, en primera persona, de aquella doble relación que consideramos con los ocho experimentos propuestos por Hume, sino que surgen porque la simpatía eleva hasta

el grado de una impresión ciertas ideas de otras personas sobre nosotros. En este caso, digamos, el orgullo o la humildad surge en nosotros porque reconocemos ideas u opiniones relativas a nuestra condición que, provenientes de otras personas, recibimos de alguna manera. Después, la simpatía opera sobre ellas, con base en la simpatía y la contigüidad, y hace surgir a las pasiones que consideramos. De este complejo proceso consideramos que se sigue, en primer lugar, que la simpatía no es ella misma una pasión o emoción, sino una condición de las pasiones (v. Kemp Smith, 189; Merrill, 266; Brand 1992, 66-7 y Fate Norton 2000, 155).⁴⁴ En efecto, la simpatía tiene la función de comunicar sentimientos (cf. T 2.1.11.19 SB 324) y en esto su función es análoga a la creencia, la cual también comunica parte del sentimiento distintivo asociado con una impresión y eleva así la fuerza y vivacidad de una cierta idea (cf. T 2.1.11.7 SB 319; Kemp Smith, 222, 377 y Merrill, 267). Observamos así, una vez más, cómo las mismas determinaciones de la mente dan lugar a los diversos estados mentales. El procedimiento por el cual la simpatía logra esto requiere considerar, bajo esta perspectiva, a la semejanza y, ahora sí, a la contigüidad.

La idea que un agente eleva al grado de pasión es, en un primer momento, una pasión de alguien más. Alguien se alegra o lamenta por uno, o manifiesta burla, desdén o admiración, etc., y para uno, inmediatamente, estas ideas son convertidas en la impresión que representan. Lo que antes era considerado meramente como una idea, una representación de un cierto estado de cosas, rápidamente es transformado en nuestra mente en un sentimiento y éste, *esencialmente*, se convierte entonces en la pasión que ahora experimentamos. Hume detalla así este procedimiento:

...es evidente que la naturaleza ha preservado una gran semejanza entre todas las distintas criaturas humanas, y que nos es imposible advertir en los demás una pasión o principio cuyo paralelo no encontremos en nosotros mismos. Lo mismo ocurre en la fábrica de la mente que en la del cuerpo (T 2.1.11.5 SB 318).

La «fábrica de la mente», por medio de la semejanza, relaciona formalmente a unas ideas con otras, a unos sentimientos con otros y, vemos ahora, también a otras personas, sus pasiones o principios: “El hombre de buen natural se encuentra al instante del mismo humor que sus compañeros... Un semblante jovial produce complacencia y serenidad en mi mente, mientras que

⁴⁴ Merrell nos deja ver que la simpatía, operando sobre las ideas, realiza una cierta operación cognitiva: “Por medio de la operación de la simpatía, las opiniones que otros tienen de mí (así sean de aprobación o de desaprobación, de elogio o de culpa) son transformadas de una mera idea (mi creencia) a una impresión que posee tal fuerza y vivacidad “como la pasión misma” (Merrill, 266).

otro enfadado y triste me infunde un repentino desaliento” (T 2.1.11.2 SB 317; cf. Stroud, 286).⁴⁵ Cuando tales sentimientos o pasiones se refieren a mí, entonces no sólo se comunica a mí el sentimiento o pasión en cuestión, sino que correspondientemente con su efecto emotivo produce alguna pasión. Así, si alguien admira algo que he hecho, yo experimento el regocijo propio de la admiración, pero pronto, como se trata de una relación mía con algo destacado, me lleno de orgullo. En otro caso, si alguien desdeña mi trabajo, mi primera impresión será de dolor y, correlativamente con la idea de mi yo, pronto experimentaré humildad. Sin embargo, esa primera impresión de dolor puede, bajo una cierta comprensión de la circunstancia, convertirse en indignación o en enojo si cabe pensar que el desdén está injustificado, si no hay razones para ese juicio, si no se ha valorado correctamente lo que hago, etc. Parece que la operación de la simpatía no elimina a la idea original ni su carácter representativo, sólo indica cómo ha ganado en fuerza y vivacidad, de una manera consistente con la fenomenología distintiva de las pasiones y también con la interpretación cognitiva de las pasiones. Además de la simpatía, la contigüidad es otra determinación básica que sostiene a la simpatía.

Mediante dicha determinación podemos explicar, siguiendo a Hume, cómo es que influyen de manera diferente los sentimientos y las pasiones de otras personas según estén más o menos cercanas a nosotros. Dice Hume al respecto: “Los sentimientos de los demás tienen poca influencia cuando esas personas no tienen relación con nosotros; es necesaria la contigüidad para poder comunicar los sentimientos en toda su integridad” (T 2.1.11.6 SB 318). Podemos explicar así cómo experimentamos más vivamente las relaciones que tenemos con personas cercanas a nosotros, amigos, familiares, compañeros de trabajo, etc. La simpatía hace posible, podríamos decir, que tales relaciones se constituyan con su dinamismo emocional característico. Éstas son relaciones que también tienen aspectos morales y que son importantes para entender tanto al carácter personal como a las relaciones que se establecen con otros agentes. En todo este proceso influye una cierta comprensión de las circunstancias en que se encuentra el agente (incluso desde una perspectiva histórica), así como el carácter o la personalidad del mismo (cf. T 2.1.11.2 SBN 317; 2.3.1.4 SBN 401; 2.3.2.6 SBN 411; 3.1.2.4 SBN 472; 3.3.1.4-5 SBN 575; Stroud, 211).⁴⁶

⁴⁵ De acuerdo con Stroud este procedimiento de contagio de una pasión en una persona a otra persona se debe, sobre todo, a la imaginación. En efecto, la imaginación es una determinación que también opera sobre la semejanza, su componente más básico según Hume.

⁴⁶ Scheffer, en *Human Morality* ha dedicado atención a cómo el carácter individual es propenso a ciertas creencias y la relación que tiene esto sobre la adquisición de deseos (v. p.e. Scheffer, 91).

Ahora podemos hacer un primer apunte de la importancia que tiene esta explicación para la moralidad.

La simpatía, en tanto que es una determinación formal que propicia por reflejo la transmisión de unas pasiones a otros, tiene además la capacidad para movernos a reflexionar sobre nuestras acciones y reacciones desde un punto de vista general, moral, y para hacer distinciones entre lo que aprobamos o desaprobamos según estándares morales (cf. Alanen 2005, 120, v. infra este aspecto relativo al punto de vista general en §4.3, p. 132). Hume mismo, en efecto, fundamenta a la moralidad en la simpatía: “...es manifiesto *que* la simpatía es un principio muy poderoso en la naturaleza humana... y *que* origina el sentimiento moral en todas las virtudes artificiales [como la justicia, según Hume]. Partiendo de estos puntos, cabe suponer que la simpatía origina también muchas de las restantes virtudes y que sea la tendencia al bien de la humanidad lo que haga merecedora de nuestra aprobación a una cualidad mental” (T 3.3.1.10 SBN 577-578). Por ahora, me interesa establecer la tesis de que hay una relación entre el fundamento de la moralidad, la simpatía, y el complejo proceso que da lugar a las pasiones y a los distintivos sentimientos morales. Esta relación muestra que la teoría de las pasiones de Hume tiene una relación lógica con la teoría de la moralidad, porque de los elementos constitutivos de la primera se siguen algunas consecuencias relevantes para caracterizar a los elementos morales. Continuaré examinando algunas consecuencias que tiene esta idea sobre la motivación.

3. La teoría de la motivación de Hume

En el capítulo anterior identifiqué un pasaje del *Tratado* que es prácticamente un resumen de la teoría de la motivación de Hume, según el cual el impulso fundamental de la mente humana es el placer o el dolor, esto es, lo que nos permite «experimentar pasión, acción, deseo o volición» es nuestra búsqueda del placer y nuestro rechazo del dolor (cf. T 1.3.10.2 SB 118).⁴⁷ Sin placer o sin dolor cesa nuestro pensamiento y sensibilidad, pero en la medida en que *nos parecen* «probables o improbables, ciertos o inciertos, a nuestro alcance o no, tienen *como efecto* los movimientos mentales de acercamiento y evitación, así como las sensaciones de deseo y aversión o las pasiones de tristeza, alegría, esperanza o miedo» (cf. T 3.3.1.2 SB 574). Sin placer o dolor nuestra vida mental carecería de relieve y textura. La operación de estos dos grupos de elementos constituye la condición esencial de las ideas de Hume sobre la motivación. En este capítulo me propongo estudiar y clarificar los aspectos que he enfatizado en esta paráfrasis: Primero, la naturaleza de la percepción que hace que el placer o el dolor *nos parezcan* probables o improbables, ciertos o inciertos, etc. Segundo, cómo dicha percepción se relaciona con otros estados mentales, como las pasiones o los deseos, para dar lugar a motivos específicos para pensar o actuar. Me concentraré en establecer lo más claramente posible la posición de Hume al respecto. Después de esto, caracterizaré a estos elementos como una teoría de la motivación en la que la razón desempeña un papel instrumental al servicio de las pasiones o los deseos porque creo que es la forma más afín a las ideas de Hume, a pesar de ciertos pasajes donde parece descartar que la razón desempeñe papel alguno en la motivación.

Con todo, dicha teoría de la motivación ha recibido fuertes críticas de Christine Korsgaard (Korsgaard 1996), las cuales examino. Respondo a sus objeciones desarrollando el modelo subhumeano de la motivación de Bernard Williams (Williams 1993). La propia Korsgaard desarrolla sus críticas a la teoría de la motivación mencionada considerando las ideas de Hume y Williams,

⁴⁷ En el pasaje referido dice Hume: “En la mente humana hay impresa una percepción del dolor y el placer, resorte capital y principio motor de todas sus acciones”. Como lo explica Nuyen, cuando una idea se asocia con una impresión, la idea adquiere parte de las cualidades de la impresión, especialmente su viveza. El contexto del pasaje, “La influencia de la creencia”, se refiere a la importancia de la creencia para las acciones, de modo que la idea relativa al placer y al dolor, y a la transferencia de la viveza de las impresiones, es, como dice Nuyen, “el trabajo de la naturaleza que nos permite anticipar peligros y evitarlos” (cf. Nuyen, 376-377). Me parece que, además, permite anticipar placeres. Creo que en este sentido hay que entender la idea de que los estados mentales son despertados por el placer y el dolor, por su importancia para explicar a las acciones.

por lo que creo que es viable responder a sus objeciones examinando con cuidado ciertas ideas de Williams sobre el papel de la razón. En todo este tratamiento me concentré en la siguiente pregunta: ¿cómo respondería un agente humeano a una razón práctica motivadora presente a él o ella?

La respuesta de Korsgaard es que un agente humeano (e incluso uno *williamseano*) falla en responder adecuadamente a dicha razón práctica. Según ella, un agente, en tanto que racional, tendría que responder adecuadamente a dicha razón práctica actuando o pensando según la guía de dicha razón. Sostengo que al respecto Williams ofrece una posición plausible para defender el modelo humeano de la motivación. Su respuesta a la objeción de Korsgaard es que un agente puede responder a tal razón práctica si ésta se relaciona de manera adecuada con algún elemento antecedente del *conjunto motivacional subjetivo del agente*. Ésta es la tesis principal del internismo de Williams, y creo que es una elaboración más clara y sencilla de ciertas ideas de Hume sobre la motivación. La posición final que defiendo es que la teoría de la motivación de Hume es plausible, precisando cuál podría ser el papel de la razón y mostrando que sólo puede tener una función instrumental. También defiendo que se trata de una teoría de la *deliberación* y no de la *razón* práctica.⁴⁸

Quizá el reto más importante que enfrenta una teoría de la deliberación es que no parece que pueda enfrentar correctamente cuestiones relativas a cómo caracterizar la racionalidad o la deseabilidad de los fines de las acciones. Si la razón sólo tiene un papel instrumental, entonces no parece que pueda ofrecer guía sobre los fines y la elección de éstos quedaría sujeta a factores que

⁴⁸ Sigo en esto la distinción que hace Williams (1981) entre el “razonamiento práctico” y la deliberación. Según Williams, las posibilidades de la deliberación son mucho más amplias que las del razonamiento práctico. Éste último lo caracteriza como aquel razonamiento que “lleva a la conclusión de que uno tiene razones para ϕ porque ϕ podría ser la manera más conveniente, económica, placentera, etc., de satisfacer algún elemento de S [el conjunto motivacional subjetivo de un agente dado], y esto, por supuesto, está controlado por otros elementos de S , aunque no necesariamente de manera clara o determinada” (p. 134). El razonamiento práctico queda así sujeto a las relaciones *derivativas* que hay entre los elementos ya presentes en el S del sujeto y la conclusión a la que se quiere llegar. La deliberación es más amplia porque no parece que se sujete a dichas relaciones derivativas. Por ejemplo, en el caso en que se presente un “conflicto irresoluble” entre elementos, mediante la deliberación un agente elige “a cuál se le da más peso”. Williams enfatiza que esto no implica “que haya un bien del cual proporcionan cantidades distintas”, lo que parece querer decir que en la deliberación no se pondera una solución a dichos conflictos comparando resultados o beneficios o alguna forma de utilidad (cf. p. 134). Precisamente porque se trata de conflictos irresolubles. Me parece que esto equivale a sostener que la deliberación no está necesariamente dominada por otros elementos del S , por lo que por medio de ella “un agente puede darse cuenta de que tiene razones para hacer algo que no sabía que tenía razones para hacer en absoluto” (p. 135). En efecto, si un agente quiere pasar una tarde entretenida, puede llegar a la decisión de que desea probar algo nuevo, por ejemplo, dar un paseo en bicicleta, cosa que no había considerado antes de ninguna manera.

no podrían caracterizarse como racionales para cualquier otro agente o para el mismo agente en circunstancias diferentes. Para responder esta cuestión examino una primera versión de la teoría del observador ideal. Hume no parece ofrecer respaldo a esta teoría cuando se ocupa, en general, de la motivación, pero sí cuando trata a los juicios morales. A pesar de esto, Michael Smith (Smith 1994) ofrece un planteamiento que reúne elementos compatibles con las ideas de Hume y, por lo menos, muestra que con base en dichos elementos puede ofrecerse una versión de la teoría del observador ideal como respuesta al reto mencionado en este párrafo. Presentaré estas ideas de Smith en la sección 3.5, más adelante, y cómo creo que pueden reforzar al modelo humeano de la motivación, a pesar de que en realidad es una teoría de Smith y que, en todo caso, comparte algunas ideas básicas con Hume.

Con estos antecedentes examino, en la misma sección mencionada, a los motivos morales. Mi examen de estos aspectos quedará, además, circunscrito a los pensamientos y acciones realizados *voluntariamente* por un agente, con el propósito o la intención de alcanzar un cierto fin.⁴⁹ Esto presupone que los agentes generan ciertos estados mentales relativos a los fines que se proponen alcanzar, pero si los motivos básicos están relacionados con el placer o el dolor, y parece que ésta es una relación *causal* antes que *conceptual*, podemos preguntarnos por la capacidad que tienen los agentes para formar o reflexionar sobre sus propósitos o intenciones para actuar. Creo que este planteamiento es otro caso de la pregunta de Korsgaard: ¿cómo responde un agente a una razón práctica disponible para él o ella? La peculiaridad ahora es que tal razón práctica se referiría a algún asunto moral. De modo que responderé a esta cuestión elaborando sobre los elementos estudiados hasta este punto. Además, en este contexto Hume ofrece algún respaldo a la interpretación del observador ideal sobre los juicios morales, de modo que pertinentemente puede agregarse este recurso.

3.1. Razones, pasiones y deseos en la formación de motivos para actuar

Quizá la dificultad más importante para el estudioso de las ideas de Hume sea establecer lo más limpiamente posible su posición. La primera dificultad, me parece, es determinar cómo se relacionan los elementos que nos parecen relevantes para explicar a los motivos para actuar, esto

⁴⁹ Con esto me refiero a la distinción entre acciones intencionales y acciones no intencionales. Las teorías filosóficas de la acción y de la motivación suelen interesarse por la acción intencional, que presupone deliberación. Las acciones no intencionales comprenden acciones como la respiración, la digestión, pero también los actos reflejos o los actos fallidos detectados por Freud.

es, razones, por un lado, y pasiones o deseos, por otro lado. En ocasiones él escribe como si no hubiera ninguna relación entre ambos grupos de facultades y cada una operará en regiones o mundos completamente diferentes. El siguiente pasaje es un ejemplo de esto:

[R1] El entendimiento se ejerce de dos formas diferentes: en cuanto que juzga por demostración, o por probabilidad; esto es, en cuanto considera las relaciones abstractas de nuestras ideas, o aquellas otras relaciones de objetos que sólo la experiencia nos proporciona información. Creo que difícilmente podrá afirmarse que la primera especie de razonamiento es por sí sola la causa de una acción. Dado que su ámbito apropiado es el mundo de las ideas, mientras que la voluntad nos sitúa siempre en el de la realidad, la demostración y la volición parecen por ello destruirse mutuamente por completo (T 2.3.3.2 SB 413).

Hume señala aquí que hay una brecha abismal entre el mundo de la razón, el mundo de las ideas, y el mundo de la voluntad o de la acción. Según él, la razón opera en el primer mundo y no puede causar nada en el segundo mundo. Las operaciones de la razón en el primer mundo son lógicas, esto es, demostraciones lógicas y relaciones entre ideas que, a lo sumo, equivalen a darle tratamiento lógico a la información originada en la experiencia. Por contraste, el mundo de la acción es el mundo de la realidad, de los objetos naturales espacio-temporales. La brecha entre ambos mundos es así tanto categorial como conceptual (cf. Platts 1988, 197). Ahora bien, el siguiente pasaje parece decir que no hay tal brecha, sino que existen al menos dos vías de comunicación entre la razón y las pasiones:

[R2] ...las pasiones podrán ser contrarias a la razón solamente en cuanto que acompañadas de algún juicio u opinión. De acuerdo con este principio... solamente en dos sentidos puede decirse que una afección es irrazonable. En primer lugar, cuando una pasión, sea la esperanza o el miedo, la tristeza o la alegría, la desesperación o la confianza, está basada en la suposición de la existencia de objetos que en realidad no existen. En segundo lugar, cuando al poner en acto alguna pasión elegimos medios insuficientes para el fin previsto, y nos engañamos nosotros mismos en nuestro juicio de causas y efectos. Si una pasión no está fundada en falsos supuestos, ni elige medios insuficientes para cumplir su fin, el entendimiento no puede ni justificarla ni condenarla (T 2.3.3.6 SB 416).

A primera vista, parece entonces que R1 y R2 son entre sí inconsistentes.⁵⁰ Según R2, la comunicación entre la razón y las pasiones es posible porque *algo* en las pasiones es propio del mundo de las ideas, el primer mundo indicado en R1. Este pensamiento es consistente con la

⁵⁰ Sophie Botros nos advierte que con Hume “tenemos que ser altamente selectivos en nuestro uso de las citas” (cf. Botros, p. 11). Su preocupación al hacer esta advertencia es que la teoría de la motivación de Hume parece encerrar una contradicción, precisamente entre las ideas de los pasajes que presentado como R1 y R2. Por una parte, Hume sostiene que “la razón no puede *nunca* producir una acción o dar origen a la volición” (T 2.3.3.4 SB 414-415, énfasis añadido) y “la razón es perfectamente inerte y no puede *nunca* evitar o producir una acción o afección” (T 3.1.1.8 SB 458, énfasis añadido, traducción mía). Estas tesis resumen su posición escéptica sobre la racionalidad práctica. Pero, casi en seguida, Hume parece conceder que la razón sí desempeña un papel instrumental, específicamente, cuando el fin de nuestras acciones «está basado en la suposición de la existencia de objetos que en realidad no existen» o «cuando actuamos de acuerdo con alguna pasión elegimos medios insuficientes para el fin previsto, y nos engañamos nosotros mismos en nuestro juicio de causas y efectos» (T 2.3.3.6 SB 416). De modo que tenemos prácticamente una contradicción. Botros, con todo, defiende una interpretación instrumental de las ideas de Hume (cf. Botros, 11 e infra p. 91).

interpretación cognitiva de las pasiones, según la cual, como vimos en el capítulo anterior, las pasiones están también constituidas por elementos cognitivas porque tienen al menos una idea, esto es, las pasiones son *acerca de* algo. De acuerdo con este pasaje, la percepción de ese algo tiene relevancia instrumental para la acción porque podría tratarse de un objeto que no existe y no podríamos alcanzar o porque carecemos de los medios adecuados para alcanzarlo. Hume dice que una pasión es «irrazonable» si no recibe ayuda de la razón cuando estas dos carencias instrumentales no son percibidas como tales por el agente y, de alguna manera, persiste en el esfuerzo de alcanzar el objeto de esas pasiones. Por contraste con esta idea, una pasión sería razonable si los juicios que la acompañan dan lugar a «suposiciones» correctas, esto es, cogniciones correctas respecto del objeto que pretendemos alcanzar o sobre los medios de que disponemos para alcanzarlo. Sin embargo, las ideas de la razón no son, en modo alguno, *razonables* respecto de ninguna preferencia, deseabilidad o acción. Según Hume, las ideas de la razón son motivacionalmente inertes y la deseabilidad de un bien o un curso de acción depende del modo en que las pasiones lo tengan como motivo. El único valor práctico de las ideas de la razón, por ejemplo cuando se trata de creencias sobre hechos, es instrumental, respecto de la mejor forma de alcanzar un fin o respecto de la existencia de dicho fin, pero a la razón le es indiferente el valor mismo del fin en cuestión. Así, respecto de la razón Hume escribe:

[R3] No es contrario a la razón el preferir la destrucción del mundo entero a tener un rasguño en mi dedo. No es contrario a la razón que yo prefiera mi ruina total con tal de evitar el menor sufrimiento a un *indio* o a cualquier persona totalmente desconocida. Tampoco es contrario a la razón el preferir un bien pequeño, aunque lo reconozca menor, a otro mayor, y tener una afección más ardiente por el primero que por el segundo (T 2.3.3.6 SB 416).

En consecuencia, la deseabilidad y la motivación parecen tener como centro la idea de uno mismo o el yo. Según vimos, este es el rasgo distintivo más importante de las pasiones indirectas y lo que Hume parece decir aquí es que la única forma en que la motivación esté acompañada de juicios sobre la deseabilidad de las acciones es que tales juicios se refieran relevantemente al bien del yo –según vimos antes para Hume el bien y el placer podrían coincidir, tanto fenomenológica como conceptualmente (T 2.1.1.2 SB 276, T 2.3.3.10 SB 418; Karlsson, 263, Brand, 24).⁵¹ Bajo esta consideración, en efecto, cabe rechazar la destrucción del mundo y preferir, si así lo requiere el beneficio del yo, un rasguño en mi dedo. En general, sólo bajo la perspectiva del beneficio de uno

⁵¹ Las pasiones calmadas, según el primer pasaje del *Tratado* citado, comprenden también al sentimiento de belleza y deformidad, que son sentimientos estéticos. En el segundo pasaje, Hume identifica como “fuerza del espíritu” [*strength of mind*] al efecto de las pasiones calmadas para prevalecer sobre las pasiones violentas. Más abajo presento como a este pasaje [P3] y lo discuto con mayor detalle.

mismo buscamos los mayores bienes y rechazamos los menores. Ésta es otra conexión entre la teoría de la motivación y la teoría de las pasiones que no suele reconocerse por estudiosos de Hume.⁵² Los juicios pueden ser razonables o irrazonables, verdaderos o falsos, pero «no dan ningún veredicto», como dice Kemp Smith (cf. 196-197), esto es, por sí mismos no nos dicen qué hacer, qué es lo deseable o lo mejor para nosotros. Los juicios sólo adquieren valor práctico o ganan en deseabilidad y valores prudenciales cuando sirven a la idea de uno mismo. En consecuencia, Hume sostiene que “la razón no puede nunca producir una acción o dar origen a la volición” (T 2.3.3.4 SB 414-415). En el terreno de la moralidad reitera esta afirmación: “la razón consiste en algo perfectamente inerte y [...] nunca puede evitar o producir una acción afección” (T 3.1.1.8 SB 458). Ahora bien, tomados al pie de la letra los pasajes R1 y R2 pueden parecer entre sí inconsistentes, pero manteniendo a la vista R2 y R3 me parece que puede ofrecerse una explicación persuasiva del modo en que la razón se relaciona con las pasiones.

La razón, en efecto, nunca puede por sí misma evitar o producir una acción. Aun cuando podamos *suponer* que, por algún acto de la mente, pueda ocuparse de examinar el valor instrumental o la deseabilidad de ciertas ideas relativas a hechos, R2 y R3 acotarían que tal examen se haría *instrumentalmente* en relación con algún fin señalado por las pasiones o en relación con la idea del yo, también constitutiva de las pasiones indirectas –recordemos que para Hume todo acto de la mente, incluyendo los actos de la razón, tiene por motor o causa al placer o el dolor, es decir, un estado no racional. R2, interpretado de acuerdo con la teoría cognitiva de las pasiones expuesta en el capítulo anterior, parece suficiente para indicar que no hay inconsistencia con R1, pues las pasiones están constituidas por ideas y, según permite suponer R3, en particular por la idea del yo. Así que la razón se ocuparía precisamente de tales ideas. En resumen, parece

⁵² Wiggins, por ejemplo, cree que este pasaje prueba que para Hume la racionalidad es una creación social, civilizadora, originada directamente en la naturaleza humana (cf. Wiggins 2006, 79). Creo que la explicación que ofrezco yo, que constituye una forma de egoísmo ético, parece más consistente con las ideas de Hume sobre las pasiones y el contexto general del examen de la motivación. Platts, por su parte, considera que este pasaje es una respuesta a la brecha que hay entre estados cognitivos y no-cognitivos, a la que él llama “la tesis aislacionista”. De acuerdo con la respuesta de Hume es una tesis causal que mantiene la brecha pero otorga a los estados cognitivos poderes causales para despertar estados no-cognitivos (cf. Platts 1988, 195 y ss.). Stroud indica que Hume parece dejar fuera a la razón como una facultad de la mente, al parecer porque ella no interviene directamente en los procesos deliberativos y porque Hume parece considerar sólo a los objetos de la razón, es decir, el conjunto de proposiciones que acompañan a las pasiones (cf. Stroud, 235-6). Baillie enfatiza que el pasaje es suficiente para descartar que las pasiones carezcan de «contenido intencional», lo cual coincide con el punto que quiero enfatizar con la interpretación cognitiva de las pasiones, pero no hace, hasta donde puedo apreciar, ninguna conexión con las pasiones indirectas o con la idea de uno mismo o el yo (cf. Baillie, 90-3). Por último, Karlsson se limita, con su apreciación de este pasaje, a señalar que el *locus* de la motivación lo constituyen elementos conativos como necesitar, desear o procurar [*caring*], los cuales de alguna manera parecen trabajar junto con la razón, pero subordinadamente (cf. Karlsson, 237).

que la explicación de la motivación de Hume sostendría que la razón sirve instrumentalmente a las pasiones considerando las ideas que son constitutivas de aquellas. Este papel de la razón parece que puede realizarse en dos direcciones: sirviendo hacia *adentro* de la pasión mediante la consideración de la idea del yo o sirviendo hacia *afuera* mediante la consideración de los objetos a los que la pasión aspira, así sean los medios para alcanzarlo o el objeto mismo. No parece representar ningún problema para esta explicación la cuestión de cómo de la elección de ciertas acciones puede derivarse la aniquilación del mundo y, por consiguiente, de uno mismo. La indeseabilidad de esta posibilidad es manifiesta a la mente y está suficientemente cargada de elementos emotivos y racionales como para despertar su rechazo, mediante actos mentales o las acciones apropiadas al caso. Con todo, en otros pasajes sobre la motivación y las pasiones Hume escribe de manera que nos conduce a pensar que la relación entre la razón y las pasiones no parece conceptual, sino contingente de manera causal, esto es, que la razón no interviene instrumentalmente en la motivación por la consideración sobre el contenido cognitivo de las pasiones –las ideas que las constituyen–, sino que interviene de alguna manera todavía misteriosa cuando las pasiones parecen insuficientes para determinar los medios de satisfacerse. Consideremos, al respecto:

[P1] Es obvio que cuando esperamos de algún objeto dolor o placer, sentimos una emoción consiguiente de aversión o inclinación, y somos llevados a evitar o aceptar aquello que nos proporciona dicho desagrado o satisfacción. Igualmente es obvio que esta emoción no se limita a esto, sino que, haciéndonos volver la vista en todas direcciones, percibe qué objetos están conectados con el original mediante la relación de causa y efecto. Aquí, pues, tiene lugar el razonamiento para descubrir esta relación, y, según varíe nuestro razonamiento, recibirán nuestras acciones una subsiguiente variación. Pero en este caso es evidente que el impulso no surge de la razón, sino que es únicamente dirigido por ella. De donde surge la aversión o inclinación hacia un objeto es de la expectativa de dolor o placer (T 2.3.3.3 SB 414).

«Cuando esperamos placer o dolor de algún objeto», desde el punto de vista de la pasión (si podemos decirlo así), tiene lugar una operación que ya parece cognitiva pues contiene, al decir de Karlsson, «un juicio sobre la expectativa de dolor o placer» (cf. Karlsson, 246). Esto es, la pasión misma parece tener la facultad o capacidad de determinar si un objeto dado puede producir placer o dolor.⁵³ En un momento inmediatamente posterior, la percepción de la pasión parece elaborar

⁵³ Recordemos como las creencias sobre el futuro se derivan para Hume por completo del hábito: “...podemos observar que la suposición de *que el futuro es semejante al pasado* no está basada en argumentos de ningún tipo, sino que se deriva totalmente del hábito, por el cual nos vemos obligados a esperar para el futuro la misma serie de objetos a que estamos acostumbrados. Este hábito o determinación para transferir el pasado es pleno y perfecto; y, por consiguiente, el primer impulso de la imaginación en esta especie de razonamientos está dotado de las mismas cualidades” (T 1.3.12.9 SB 134). De este modo, puede aceptarse que las creencias caracterizadas como expectativas, que son creencias sobre el futuro, no surgen de la razón.

otra especie de juicio mediante el cual determina las conexiones del objeto original con otros que lo rodean y que, según sea el caso, es necesario disponer de alguna manera para alcanzar al objeto original o para alejarse de él, de acuerdo con el primer juicio al respecto, *i.e.*, el juicio sobre el placer o el dolor que nos depara tal objeto. En este punto, y sólo en este punto según P1, la razón es requerida. Su función parece ser la de trazar la ruta causal que va de nuestras deliberaciones hasta alcanzar la situación que anhelamos con el objeto de las mismas. Esta es una ruta *instrumental* porque mediante ella, parece decir Hume, la razón concebiría los medios para alcanzar o para alejarnos del objeto que despierta todo el proceso motivacional. Hume continúa diciendo que “según varíe nuestro razonamiento, recibirán nuestras acciones una subsiguiente variación” lo que parece equivaler a la observación de que el razonamiento causal-instrumental modifica el rumbo de la acción, por ejemplo, si nos decidimos por una vía u otra para alcanzar la situación que anticipamos tener con el objeto. También parece equivaler a decir que la pasión podría modificar su primer juicio, el que detona todo el proceso, según los dos modos considerados en R2, pero en este caso la razón sólo influye mediante el juicio causal, el segundo juicio que la pasión parece elaborar. Más importante aún parece ser que la razón misma no tiene la última palabra sobre seguir un curso de acción dado para alcanzar el objeto. Podría ser que tal curso de acción resultará demasiado costoso o que expusiera al yo a una condición precaria o, incluso, que no fuera consistente con los objetos de otras pasiones del yo.

Para Hume la razón, como hemos visto con R3, no puede de ninguna manera determinar las condiciones de deseabilidad de ninguna acción, así sea ésta intermedia o final. En cada caso, la teoría de Hume señala que es la pasión la que establece por sí misma las condiciones de deseabilidad. Además, según sabemos por el examen de las creencias sobre hechos del capítulo 1, la razón no es especialmente apta para considerar las relaciones causales. La experiencia, o la costumbre, lo hacen mejor que ella y, según parece decirnos Hume en el siguiente pasaje, la pasión dictamina tanto sobre las consideraciones instrumentales de la razón como de la experiencia:

[P2] De donde surge la aversión o inclinación hacia un objeto es de la perspectiva de dolor o placer. Y estas emociones se extienden a las causas y efectos de este objeto, tal como nos son señaladas por la razón y la experiencia. Nunca nos concerniría en lo más mínimo el saber que tales objetos son causas y tales otros son efectos, si tanto las causas como los efectos nos fueran indiferentes (T 2.3.3.3 SB 414, énfasis añadido).

Hasta donde veo, los estudiosos de Hume no suelen reparar en ese “y la experiencia” enfatizado arriba. Creo que con ello Hume diezma, a la luz de su teoría de las creencias sobre hechos, en buen grado la intervención de la razón en los procesos deliberativos, porque en muchos casos la

pasión y la experiencia serían suficientes para explicar la eficacia de nuestras deliberaciones y acciones para alcanzar lo que nos proponemos. Además, recordemos cómo para Hume la «experiencia» reúne las determinaciones naturales de la mente, esto es, los principios de asociación entre ideas, la imaginación o el hábito (cf. supra § 1.1), de modo que en la motivación también estas determinaciones intervienen ayudando instrumentalmente a las pasiones. Sea la razón o la experiencia la que asiste a las pasiones el punto de P2 es que la última palabra sobre la volición y la acción la tienen éstas. En efecto, desde el punto de vista de la razón (si se puede hablar así) es indiferente que algo sea causa o efecto, incluso si se trata de tomar medidas para actuar. La razón, como he dicho a propósito de R3, no discierne ni dictamina de acuerdo con ningún criterio de preferencia. La experiencia, por ejemplo entendida como hábito, sigue otras pautas, pero en este caso no hablaríamos de una facultad cognitiva. Las pautas del hábito son también preferencias o deseos de algún tipo que se han anquilosado en la consciencia y al menos no parecen preferencias racionales. Así, para que sea relevante distinguir entre causa y efecto o para establecer aquello que sirve mejor a la idea del yo ante una acción o una decisión la razón no ofrece «dictamen» alguno e, impulsada por la fuerza volitiva de las pasiones que empuja al placer o aleja del dolor, simplemente señala los medios para los fines de las pasiones. Si la razón establece, por ejemplo, varios cursos de acción o un curso de acción que represente un costo muy grande no es ella la que determina qué hacer, sino la pasión la que dictamina qué se hace y cuáles son los costos que pueden asumir nuestra idea del yo. Por ejemplo, como a nuestra idea del yo el deseo de preservar la salud es muy valioso, podemos someternos con pocas reservas a dolores menos costosos (en términos de placer) como visitar al dentista o no siempre beber o comer lo que más nos antoja. De acuerdo con esto, siempre que actuamos, por el motivo que sea, ese motivo será más fuerte y representará un bien preferido en algún grado para nuestro yo –de este modo, para la teoría de la motivación de Hume no hay lugar para el fenómeno de la debilidad de la voluntad, pues siempre que supuestamente surja un caso que pueda caer bajo él, siempre podremos suponer que hay un motivo de mayor peso que fue el que determinó la acción del agente. En efecto, Hume supone que la fuerza de las pasiones no siempre se experimenta fenoménicamente. Según él, muchas veces las pasiones actúan de manera «tranquila» y esta fuerza es la que, finalmente, explica a la acción:

[P3] Los hombres actúan a menudo conscientemente en contra de sus propios intereses. Por esta razón la consideración del mayor bien posible no siempre tiene influencia sobre ellos. Los hombres reprimen con frecuencia una pasión violenta al perseguir sus intereses y fines. Por tanto, no es

solamente el desagrado presente lo que les determina. [...] Lo que llamamos fuerza de espíritu [*strength of mind*] implica el predominio de las pasiones apacibles [*calm passions*]... (T 2.3.4.10 SB 418).

Sobre este pasaje, Audi ha señalado que en él Hume “afirma la posibilidad de un tipo de la debilidad de la voluntad” y más adelante explica que cuando habla de fuerza de espíritu –*strength of mind*– “él claramente no está hablando de fuerza intelectual, y podemos suponer razonablemente que quiere decir algo como fuerza de voluntad” (cf. Audi, 48-49). Estoy de acuerdo con Audi en esto, pero creo que su idea de que «la fuerza de la voluntad implica el predominio de las pasiones apacibles» indica que él no sostendría que la debilidad de la voluntad es un hecho psicológico genuino, porque en cada caso, aun cuando no tengamos registro consciente de ello, siempre estaríamos actuando de acuerdo con un deseo inmediato y con la fuerza suficiente para motivar pensamientos o acciones determinadas.

De acuerdo con Hume las pasiones calmadas «determinan con frecuencia a la voluntad», pero lo hacen de manera imperceptible y esto propicia que «quienes no examinan los objetos de un modo estrictamente filosófico las confunden con la propia razón». Las pasiones calmadas son caracterizadas, en su teoría de la motivación, del siguiente modo:

[P4] ...existen ciertos tranquilos deseos y tendencias que, a pesar de ser verdaderas pasiones, producen poca emoción en el alma, y son más conocidos por sus efectos que por su sentimiento o emoción inmediata. [...] Cuando alguna de estas pasiones está en calma, sin ocasionar desorden en el alma, muy fácilmente se confunde con las determinaciones de la razón, por suponer que procede de la misma facultad que juzga de la verdad y la falsedad (T 2.3.4.8 SB 417).

Según Baillie las pasiones calmadas son una anomalía para el modelo de la mente de Hume, porque la distinción básica entre impresiones e ideas descansa en la diferencia fenomenológica de la fuerza respectiva de cada uno de tales materiales básicos (cf. Baillie, 22), de modo que correspondería a las pasiones, un tipo de impresión, caracterizarse por la viveza de sus manifestaciones. Sin embargo, parece que Hume tiene que postular a las pasiones calmadas en vista del papel pasivo e inerte que ha asignado a la razón de cara a la acción (v. arriba R1 y R3 y Mackie 1980, 47).⁵⁴ Ante esto, tiene que haber una fuerza no-cognitiva que pueda generar a los motivos que determinan a la voluntad.

⁵⁴ Mackie argumenta que Hume tiene que postular a las pasiones calmadas bajo una interpretación claramente cognitiva de las pasiones. Según él, Hume ha llegado a la idea de que las creencias nos mueven a la acción solo si son relevantes para la satisfacción de una pasión, y que razonar, sea demostrativa o probablemente, afecta nuestras acciones sólo en la medida en que produce creencias también relevantes, pero como la razón no puede producir directamente a la acción, entonces tiene que haber algo, en el orden de las pasiones, que produzca la acción (cf. Mackie 1980, 47).

Con esta estrategia argumentativa Hume ha llegado a una idea muy plausible sobre la vida mental de los agentes. Con su defensa de las pasiones calmadas Hume refuerza su idea de que no tenemos acceso ni percepción de todos nuestros estados mentales. De acuerdo con él muchos de los procesos mentales y de las determinaciones de nuestra conducta y pensamiento tienen lugar de manera que no tenemos conocimiento directo de ellos sino que tales procesos son conocidos indirectamente «por sus efectos y no por su sentimiento o emoción inmediata». Al respecto, llama la atención que Félix Duque, quizá el traductor más autorizado de Hume al español, haya elegido caracterizar este rasgo oscuro de los procesos mentales como una transición ‘inconsciente’, pero al hacerlo así incurre en un anacronismo (Hume no disponía de este término). El pasaje en cuestión, según su traducción, dice: “Cuando nos limitamos a seguir la determinación habitual de la mente efectuamos la transición inconscientemente [*without any reflection*], sin interponer ni un instante entre la visión de un objeto y la creencia en su acompañante habitual” (T 1.3.12.7 SB 133). Este anacronismo no altera el sentido de la idea de Hume y parece quizá aludir a su contribución conceptual a la manera en que hoy en día, comúnmente, solemos hablar de la mente. En efecto, en el pasaje en cuestión la idea es que la costumbre forma a la idea de causalidad, sin necesidad de deliberación alguna, pues tal idea «surge sin reflexión alguna», esto es, sin la participación de la razón. Podemos apreciar así otro aspecto importante de su intuición sobre el *inconsciente*: las operaciones de la mente que determinan a la voluntad son inaccesibles a la reflexión y por tanto, no son y no pueden ser objeto de la razón. Nuestras voliciones obedecen a determinaciones naturales que escapan a nuestro control y conocimiento directo y, no obstante, son causalmente eficaces para producir pensamientos y voliciones. De otro modo: el origen de los motivos por los que actuamos es oscuro y está enraizado profundamente en la subjetividad de cada uno, fuera del alcance de la razón. Parece, además, que el argumento de Hume sobre las pasiones calmadas nos aporta elementos cruciales para el tratamiento de los motivos para actuar, en la línea de una interpretación no-cognitiva de dichos motivos.

El tratamiento que hace Hume de las pasiones calmadas apunta a una teoría fuertemente no-cognitiva de la motivación. Sin embargo, como hemos visto a propósito de R1, R2 y R3 y de su teoría cognitiva de las pasiones, la razón al menos puede desempeñar un papel instrumental en todo este proceso. Este papel instrumental, según sugiero, está marcado por la percepción de aquello que sirve a la idea del yo, de uno mismo en cada caso. Este último aspecto marca condiciones de deseabilidad fuertemente relativas a uno mismo. La razón es así un instrumento

del yo –una mera «esclava de las pasiones»–, pero al mismo tiempo las pasiones responden bien, en aquellos casos en que el juicio que las acompaña es manifiestamente irracional según los dos casos que Hume considera en R2, es decir, las pasiones se someten a los dictámenes de la razón sobre la verdad o falsedad sobre los fines o sobre los medios para satisfacerse: “En el momento mismo en que percibimos la falsedad de una suposición o la insuficiencia de los medios [cf. R2], nuestras pasiones se someten a nuestra razón sin oposición alguna” (T 2.3.3.7 SB 416; cf. Fate Norton 2000, I72).⁵⁵ De este modo, parece que la colaboración entre la razón y las pasiones es muy estrecha y, al menos en la esfera instrumental, la razón parece ejercer una gran influencia sobre las pasiones. Con todo, desde la perspectiva de la deseabilidad de las acciones y pensamientos estas observaciones parecen poco interesantes porque poseen muy poco alcance.

A pesar de las dificultades que pueda haber para interpretar correctamente el pensamiento de Hume, me parece que en su tratamiento del papel de la razón en la motivación ha puesto cuidado en limitarla a un papel negativo: la razón puede extinguir a las pasiones presentando la falsedad de su objeto o la insuficiencia de los medios para alcanzar un objeto y, en general, nunca puede producir una acción o dar origen a una volición porque es perfectamente inerte respecto de la acción o de las voliciones. Así, la razón no parece tener ningún papel positivo en la motivación, excepto establecer *algunas veces* los medios para alcanzar un cierto fin señalado por las pasiones (cf. P2 para este ‘algunas veces’, porque, como indica Hume, otras veces la experiencia sola basta a la pasión para satisfacerse). De acuerdo con esto, las pasiones solas, y sólo ellas, influyen sobre la voluntad y determinan qué acciones realizamos o los pensamientos correspondientes. La limitación sobre la deseabilidad de tales acciones o pensamientos se aprecia si consideramos que en su génesis no intervienen procesos reflexivos, porque o bien obedecen a los objetos más profundos de las pasiones (cf. antes cap. 2. p. 49) o bien no son accesibles a ellos por pertenecer al orden de las pasiones calmadas, esto es, a los estados no-cognitivos que dirigen muchas veces nuestro pensamiento y nuestra acción (cabe suponer que estos dos órdenes no-cognitivos de las pasiones se traslapan de alguna manera, pero no consideraré por ahora esta dimensión explicativa del modelo humeano de la mente). Con base en esto, y solo en esto, ¿cómo se establece la

⁵⁵ Comentando este pasaje, Fate Norton indica que la razón sirve como los ojos y oídos de las pasiones. Así, aunque no se puede oponer directamente a las pasiones, tiene gran influencia sobre ellas. Continúa diciendo que nuestras pasiones, según hemos descubierto, dependen de juicios o creencias, sobre las cuales la razón tiene un papel importante en determinarlas (cf. Fate Norton 2000, I72). Mi exposición sigue estas ideas, pero creo que concedo más peso al argumento de Hume sobre las pasiones calmadas y sus implicaciones sobre la mente de lo que parece conceder Fate Norton en su exposición de la teoría de la motivación de Hume.

deseabilidad de un pensamiento o una acción sin la ayuda de la razón? Parece que la única opción disponible para Hume motivación es la idea de uno mismo o del yo. Sin embargo, esta idea está construida con los mismos materiales que escapan a la consideración reflexiva o a la razón. Como hemos visto, el propio yo parece un producto de los estados conativos, de las pasiones en particular, y de los procesos de largo plazo donde se consolida el objeto formal de las pasiones a partir de impresiones particulares del objeto profundo (cf. antes cap. 2). De manera que nuestra propia idea del yo no es reflexiva, esto es, en su formación no interviene la razón, tal como no interviene en los procesos que dan lugar a los motivos por los que actuamos o pensamos. Si la deseabilidad de un pensamiento o de una cierta acción es fuertemente relativa a la idea que cada agente tiene de sí mismo, entonces tenemos que contar con criterios de deseabilidad muy particulares, muy individuales o subjetivos. En consecuencia, no parece que la deseabilidad, según este modelo, pueda entenderse bajo criterios racionales ni tampoco objetivos. Bajo esta caracterización de la deseabilidad, parece que los agentes humanos sólo aceptarían una razón para actuar como propia si a ésta la acompaña un deseo o pasión propias, incluso si de alguna manera pudiera aceptarse que de alguna manera el agente *crea* tal razón. Las razones morales nos ofrecen, quizá, el caso más importante al respecto.

3.2. La teoría instrumentalista de la motivación de Hume

Hasta aquí, la explicación de Hume sobre la motivación se apoya en las siguientes afirmaciones:

1. Las pasiones y la razón pertenecen a categorías distintas, las primeras operan en el mundo de la voluntad y la acción, mientras que la segunda lo hace en el mundo de las ideas o del pensamiento.
2. Las pasiones hacen surgir los motivos por los que actuamos o pensamos, de acuerdo con nuestras determinaciones naturales de buscar el placer y alejarnos del dolor y bajo una cierta operación cognitiva por medio de la cual establecen alguna expectativa respecto del placer o dolor que algún objeto guarda para ellas.
3. Las pasiones son cognitivas y están constituidas por ideas. Sobre éstas opera la razón de manera instrumental, sirviendo a la pasión de dos únicas maneras: señalando si se ha fijado en un objeto que no existe o si se han elegido medios insuficientes para alcanzarlo.
4. También la experiencia y el hábito (además de la razón) sirven instrumentalmente a la pasión.
5. Siempre actuamos motivados por alguna pasión, así sea una pasión calmada. Bajo este tipo de pasión nuestros motivos no son percibidos fenoménicamente (porque carecen de la «fuerza» distintiva de las impresiones). En consecuencia,

6. No hay lugar en esta explicación para lo que suele llamarse ‘debilidad de la voluntad’, porque en cada caso los agentes actúan siguiendo el deseo más fuerte, aunque no tengamos registro en la consciencia de él.⁵⁶
7. Las operaciones de la razón carecen de valor *práctico* (la razón sólo hace operaciones relativas a la verdad de creencias, juicios o enunciados y a la corrección de razonamientos). Su único valor es teórico e instrumental como lo delimita (3), pero es suficiente para influir sobre la pasión, por ejemplo, extinguiéndola si señala que su objeto no existe.
8. No siempre tenemos acceso a nuestros estados mentales, o percepción de ellos, incluso de aquellos que determinan los motivos que influyen sobre la voluntad o el pensamiento.
9. Desde un punto de vista individual, sólo mediante una consideración relativa a la idea de nuestro propio yo podemos evaluar la deseabilidad de un motivo para actuar o para pensar.
10. La idea del yo no es cognitiva ni se ha formado reflexivamente. Su origen está asociado con y quizá determinado por el origen mismo del objeto profundo constitutivo de la pasión.

Ahora bien, para que estas ideas puedan articularse adecuadamente como una teoría de la motivación que deja a la razón un papel instrumental creo que es necesario que con ellas podamos además respaldar estas dos afirmaciones: la primera, que la agencia es causalmente eficaz aunque deliberativamente mínima y, la segunda, que los criterios de deseabilidad son subjetivos. En efecto, los agentes humanos parecen agentes causales eficaces, aunque las condiciones de deliberación que podemos atribuirles parecen, por lo menos, mínimas desde un punto de vista racional. Sin embargo, también parece justo reconocer que las ideas del sumario anterior no permiten entender en qué podría consistir un punto de vista racional sobre la acción y el pensamiento, es decir, no parecen dar lugar a algo como el punto de vista de la «razón práctica», sea cual sea éste. En efecto, exigirle a la teoría de Hume una respuesta racional a los problemas prácticos, como el de deseabilidad de un fin o acción, parece obligarla a hablar en un lenguaje ajeno. La segunda afirmación es consistente con la primera: en estas condiciones, si en efecto la razón no interviene en la ponderación de los fines de las acciones humanas, parece que al menos podemos aceptar aquellas consideraciones instrumentales que sirven a los fines marcados por las pasiones.

⁵⁶ En un sentido, el fenómeno de la debilidad de la voluntad es “amigo” de una interpretación externista de la teoría de la motivación de Hume, según lo pone Sara Stroud (2003, 129-130). Según esta interpretación, un externista de la motivación aceptaría la tesis de que no hay ninguna conexión necesaria o interna entre un juicio práctico y la acción (o intención o motivo) (cf. S. Stroud, 126, nota 7). La teoría de Hume sería externista, en este sentido, porque niega que haya alguna conexión entre la deliberación y cualquier juicio, a menos que un deseo acompañe al juicio, pero en este caso es el deseo el que suministra el motivo para actuar.

Los defensores de una concepción instrumentalista de la razón, como parte de la teoría de la motivación de Hume, aceptan que se puede explicar *adecuadamente* la relación entre los medios para alcanzar un fin y dicho fin y que ésta es una relación racional o, por lo menos, *regulada* en alguna medida por pautas racionales. Según Sophie Botros (Botros 2006), una defensora de esta concepción aceptaría las siguientes tesis como constituyentes de tal concepción instrumentalista:

- (1) La razón *puede* influir sobre la acción –en el sentido de que una consideración racional puede ser relevante para explicar una cierta acción x, al menos de manera instrumental;
- (2) Esto no es, sin embargo, suficiente para la acción dado que se requiere de la ayuda del deseo;
- (3) El deseo es necesario para la acción;
- (4) El rol de la razón es diferente al del deseo en que ella provee información relevante respecto de qué deseos tener y cómo satisfacerlos, mientras que los deseos proveen nuestros fines;
- (5) El rol de la razón no es casual, a diferencia del deseo.

De acuerdo con Botros, Hume suscribe estas 5 tesis (cf. Botros, 11), que en conjunto respaldan una interpretación instrumentalista del papel de la razón en la motivación. Estoy de acuerdo con esto y me parece que estas cinco afirmaciones reúnen adecuadamente los requisitos para una teoría de la motivación que le da a la razón un papel instrumental.⁵⁷ La estrategia de Botros para ofrecer una interpretación instrumentalista de la teoría de la motivación de Hume, consiste en “tratar al argumento práctico de Hume como *si* comprimiera en su aparente estructura unitaria dos argumentos...” Una primera interpretación ofrece una versión ‘moderada’ del argumento y la otra una versión ‘extrema’ (cf. Botros, 1). Botros sostiene que la versión moderada es correcta y que se ajusta con las tesis presentadas arriba. La versión extrema suscribe y extrae diversas consecuencias de: “...la razón es perfectamente inerte y no puede *nunca* evitar o producir una acción o afección” (T 3.1.1.8 SB 458, énfasis añadido). Botros considera que la versión extrema

⁵⁷ Jane Hampton ofrece un conjunto distinto de tesis instrumentalistas que parece más fuerte que éste y, quizá por ello mismo, más propenso a controversia. Según ella, un defensor de esta concepción aceptaría las siguientes tesis como constituyentes de tal concepción instrumentalista: 1. Una acción es racional en la medida en que el agente cree (razonablemente) que conducirá a la obtención de un fin; y (2) Los razonamientos humanos involucran la determinación de los medios para alcanzar fines, de un modo descrito por la teoría (p.e., que el uso de la razón determina hasta qué punto una acción es un medio para un cierto fin es un uso instrumental de la razón); y (3) De ninguna manera estos fines son puestos por la razón operando no-instrumentalmente; i.e., lo que los hace nuestros fines es algo distinto a la razón (Hampton, p. 57). De acuerdo con Hampton, una concepción instrumentalista de la razón tendría que aceptar necesariamente estas tres tesis (Hampton, 58). Ella defiende que Hume no apoya tal concepción y creo que esto es correcto: Hume no parece un instrumentalista que suscriba este conjunto de tesis.

es insostenible. Para establecer esta posición, sigue la interpretación defendida por Millgram en “Was Hume a Humean” (1995).⁵⁸ Con todo, no creo que sea necesario repasar estas interpretaciones y discusiones sobre la viabilidad o inviabilidad de ellas. Millgram defiende su posición siguiendo al pie de la letra ciertos pasajes de Hume, mientras que Botros recomienda hacer una lectura selectiva de los mismos. A mí me parece que mi presentación de estos temas y la selección de estos pasajes permiten ofrecer una interpretación plausible de la teoría de la motivación de Hume como instrumentalista, de modo que sólo usaré estas ideas de Botros para darle forma a la teoría, según el listado de requisitos citado arriba.

3.3. La teoría de la motivación de Hume y la acusación de escepticismo de Korsgaard

Quizá la prueba más fuerte para la teoría de la motivación de Hume proviene de la revisión que Korsgaard ha hecho de ella. Como vimos, la pregunta que tendría que responderse, satisfactoriamente por una teoría de este tipo, es: ¿cómo responde un agente humeano a una razón para actuar disponible para él? De acuerdo con Korsgaard la respuesta a esta pregunta marcaría la plausibilidad de una teoría de la racionalidad práctica, incluso en su versión instrumental. Una teoría racional más fuerte que una instrumental no tiene problemas para responder la pregunta anterior porque su tesis principal sería, en efecto, que los agentes racionales actúan de acuerdo con los dictámenes de la razón, esto es, ponderan racionalmente los fines de una acción y su

⁵⁸ En efecto, Millgram señala que es un “asunto textual [que] el Hume del *Tratado* no es de ninguna manera un instrumentalista”, esto es, que literalmente Hume no mantiene una concepción instrumental de la razón (cf. Millgram 76). El punto de la interpretación de Millgram descansa, según él, en que no es obvio qué significa que «la razón sea esclava de las pasiones» y que, para determinar su significado, tenemos que considerar dos argumentos distintos de los que la afirmación de Hume parece la conclusión (cf. Millgram, 77). El primer argumento está en T 2.3.3.2 SB 413-414 y es una descripción de las funciones de la razón en el terreno *abstracto* de las demostraciones deductivas y probabilísticas. Hume señala, además, que la razón se relaciona con razonamientos *empíricos*, digamos en cálculos mecánicos o mercantiles. En ambos casos, en los razonamientos abstractos o en los empíricos, la razón se limita a considerar las «proporciones entre números por medio de sus operaciones distintivas». En consecuencia, concluye Hume, «la razón nunca influye sobre ninguna de nuestras acciones», *i.e.*, que no tiene influencia práctica. Millgram observa que la conclusión está explícitamente asentada y que si el argumento es correcto, entonces que la razón sea esclava de las pasiones significa precisamente esto (cf. Millgram, 76). El segundo argumento, además, refuerza esta conclusión (cf. Millgram, 77). Este argumento está en T 2.3.3.5 SB 415 (y en T 3.1.1.9 SB 458). Su primera premisa sostiene, siguiendo a Millgram, que una pasión es una existencia originaria, en el sentido de que las pasiones no representan nada. La premisa dos señala que la verdad y la falsedad [de un juicio] requiere de la representación, es decir, la verdad o falsedad requiere del acuerdo o desacuerdo entre nuestras ideas y los objetos que representan. La premisa tres parece que es la que refuerza al primer argumento, porque sostiene que la razón concierne exclusivamente con la verdad o la falsedad. De estas premisas se sigue que «a una pasión no puede oponérsele la verdad o la razón; los estados prácticos de la mente no pueden ser producidos mediante razonamientos» (cf. Millgram, 77-78). Nuevamente, Millgram sostiene que “si esto es correcto, entonces su conclusión es equivalente a sostener que no hay tal cosa como el razonamiento práctico.” Además, como ambos argumentos “convergen” (por medio de la premisa 3 del segundo argumento), Millgram sostiene que Hume no es instrumentalista (cf. Millgram, 78).

respuesta es consistente con el veredicto. Al mismo tiempo, cuando sus respuestas prácticas no siempre son consistentes con el veredicto, la teoría acepta esto como un caso de irracionalidad. Son los casos en que el agente, digamos, tiene a la vista la consideración racional sobre qué es lo correcto hacer y no lo hace. Este aspecto de una teoría de la razón práctica es crucial para Korsgaard. Según ella, una teoría de la racionalidad práctica debería dar lugar a la *verdadera irracionalidad*, es decir, a la «falla en responder apropiadamente a una razón disponible» (cf. Korsgaard 1996, 318). Ahora bien, según este requisito, para Korsgaard, la explicación de Hume sobre el papel de la razón en la motivación falla tanto como una teoría de la racionalidad práctica fuerte y también como una teoría débil de la razón instrumental, (cf. Korsgaard 1996, 314), porque, de acuerdo con ella, Hume no nos ofrece una teoría donde la razón pueda darnos motivos para actuar o señalar los fines de la acción. En este sentido, no ofrece una teoría fuerte de la racionalidad práctica. Además, tampoco ofrece una teoría en donde la razón tenga un papel instrumental, por ejemplo sirviendo a las pasiones al discernir adecuadamente entre la verdad y falsedad de una creencia y actuando sobre esta distinción. De este modo, tampoco es una teoría débil de la razón instrumental. No parece haber ninguna dificultad en aceptar el primer fallo porque el propio Hume niega que la razón pueda brindar los motivos o fines últimos de la acción. El segundo fallo, de acuerdo con Korsgaard, se debe a que la teoría de Hume no admite a la verdadera irracionalidad.

Según dice la tesis (3), presentada arriba, el papel instrumental de la razón es informar a la pasión sobre la existencia de su objeto o sobre la suficiencia de los medios para alcanzarlo. Pero para Korsgaard estos dos casos fallan en cumplir el requisito de verdadera irracionalidad. Según ella, la plausibilidad de una teoría de la razón instrumental depende de que pueda reconocer y darle lugar, precisamente, a casos de genuina irracionalidad:

...parece como si una teoría racional medios/fin debe permitir al menos una forma de verdadera irracionalidad, a saber, la falla para ser motivados por la consideración de que la acción es el medio para tu fin. Incluso el escéptico acerca de la razón práctica admite que los seres humanos pueden ser motivados por la consideración de que una acción dada es un medio para un fin deseado. Pero esto no es suficiente, para explicar este hecho, que los seres humanos pueden engancharse en razonamiento causal. Es perfectamente posible imaginar una clase de ser que podría engancharse en razonamiento causal y quien podría, por lo tanto, engancharse en razonamiento que podría señalar los medios para sus fines, pero que no estuviera motivado por él (Korsgaard 1996, 319).

La idea de Korsgaard parece ser que una teoría plausible de la razón instrumental tiene que incluir y explicar aquellos casos en donde un agente por reflexión establece los medios correctos para alcanzar un fin que tiene y que, al mismo tiempo, este proceso de reflexión lo motiva a actuar

precisamente en la dirección señalada por su razón. Pero según nos dice la tesis (3), esto no sucede la deliberación de los agentes morales porque la razón juega un papel muy restringido.

Respecto del caso en que la pasión se fija alcanzar un objeto inexistente no lo es porque “las conclusiones derivadas de premisas equivocadas no son *irracionales*”. El segundo caso, cuando elegimos de hecho medios insuficientes, tampoco le parece un caso de verdadera irracionalidad porque actuar por la creencia de que tales medios podrían ser suficientes es también equivocada (cf. Korsgaard 1996, 318). En otras palabras, éstos no son casos de verdadera irracionalidad porque el agente no actúa en contra de una razón correcta o apropiada, sino con base en una razón incorrecta o falsa. La objeción de Korsgaard, como ha señalado S. Stroud, se refiere a un principio constitutivo del pensamiento racional práctico y no exactamente al compromiso que un agente cree tener entre la verdad o corrección de sus razones y la acción correspondiente (cf. S. Stroud, 136). En efecto, no parece que un agente actúe sobre la base de una creencia falsa sabiendo que es falsa. Incluso para Hume, según Korsgaard, “en relación con sus propias creencias, la gente *nunca* actúa irracionalmente” (cf. Korsgaard 1996, 319). En efecto, el propio Hume nos dice en T 2.3.3.7 SB 416, «en cuanto percibimos la falsedad de alguna suposición o la insuficiencia de los medios, nuestras pasiones se someten sin oposición a la razón», de modo que no hay lugar en su teoría para los casos de verdadera irracionalidad requeridos por Korsgaard.

Además, según la interpretación de Korsgaard, las respuestas que la teoría de Hume ofrece a las consideraciones de la razón son meramente causales.⁵⁹ Incluso si concedemos que las pasiones son de alguna manera cognitivas, por ejemplo en tanto que aceptamos que las pasiones albergan algún prospecto sobre el placer o dolor que un cierto objeto reserva para ellas (v. tesis 2 arriba), «la única preocupación que tendríamos al respecto sería sobre las causas y efectos» (cf. Korsgaard 1996, 313). Esto es, nos interesaría saber si tal expectativa de dolor o placer se conecta

⁵⁹ Korsgaard (1996), Millgram (1995) y Hampton (1995) sostienen que Hume no ofrece ninguna noción de racionalidad práctica. De acuerdo con este punto de vista, los procesos deliberativos que están detrás de las decisiones o las acciones de los agentes suceden sin la intervención de la razón. De modo que se explican exclusivamente por la intervención causal de las facultades no-cognitivas de la mente, como los deseos. Por otra parte, Audi (2006), Karlsson (2006), Botros (2006), Williams (1981), han defendido una interpretación de la razón práctica instrumental en su lectura de la teoría de la motivación de Hume, si bien con diferencias entre ellos. El punto en común es que aceptan el papel de la razón como «esclava de las pasiones», esto es, grosso modo, que la razón interviene en la deliberación práctica aportando *creencias* relevantes sobre los fines de nuestras decisiones y acciones o sobre los medios para alcanzar tales fines y medios. Así, la razón desempeña un papel *instrumental*, mientras que corresponde a las pasiones el papel teleológico, *i.e.*, establecer los fines de nuestras decisiones y acciones. En la siguiente sección (3.3.) aportaré algunas ideas para defender una interpretación instrumental de la razón práctica.

adecuadamente con el deseo correspondiente. Aun cuando nos parezca que un agente se absorbe en un proceso deliberativo, de ponderación racional de los fines de su acción, tendríamos que admitir que actúa de acuerdo con su deseo más fuerte, sea cual sea el contenido de su deliberación y la acción que realice. En la siguiente sección responderé a estas objeciones de Korsgaard.

3.4. La respuesta al reto de la “verdadera irracionalidad” de Korsgaard: una teoría humeana de la *deliberación* práctica. Las ideas de Bernard Williams al respecto

Me parece que la respuesta más directa que puede ofrecerse al escepticismo de Korsgaard sobre la teoría de la razón práctica de Hume es señalar que él no pretende ofrecer una teoría al respecto. La idea básica de una teoría de la razón práctica es que mediante la reflexión un agente humano puede abordar la cuestión de qué debería hacer en una circunstancia dada (cf. Wallace 2014). Esto equivale a decir que el agente, con algún grado de control sobre su reflexión, enfrenta la situación particular que le demanda una decisión o una acción. Por contraste, las ideas de Hume presentadas arriba sobre la deseabilidad hablan de procesos mentales de valoración, decisión y acción sobre los que el agente no parece tener control (v. arriba p. 81 y ss). Para Hume, como vimos, la explicación de por qué no siempre actuamos de acuerdo con las demandas más fuertes de los deseos y pasiones requiere de las pasiones calmadas. Para él, éstas son uno de «dos tipos de deseos»: “o ciertos instintos implantados originalmente en nuestras naturalezas, tales como la benevolencia y el resentimiento, el amor a la vida, y la amabilidad hacia los niños; o el apetito general por el bien y la aversión al mal, considerados meramente como tales” (T 2.3.4.8 SB 417; v. arriba p. 86, lo relativo a [P4]). Con esta caracterización a la vista, podemos ver «por qué su efecto se confunde con el de la razón» al determinar a la voluntad. Por ejemplo, el dolor anticipado de sufrir un rasguño en un dedo no determina que los agentes elijan la destrucción del mundo, porque “el amor a la vida” o “el apetito general por el bien” conducen a su voluntad a soportar el dolor si éste fuera en efecto el dilema. Este efecto sobre la voluntad es prácticamente imperceptible para el agente y muchas veces sólo lo reconocería una vez que ha tenido lugar. Ahora bien, creo que puede elaborarse sobre estas ideas de Hume para responder al reto de la verdadera irracionalidad planteado por Korsgaard.

Como vimos antes, este reto consiste en dar lugar en una explicación sobre la racionalidad práctica al hecho de que un agente humano muchas veces falla en responder apropiadamente a una razón disponible. En consecuencia, Hume no falla en enfrentar el reto porque, para él, “en

relación con sus propias creencias, la gente *nunca* actúa irracionalmente” (cf. Korsgaard 1996, 319; v. arriba). Sin embargo, el uso que hace Hume de las pasiones calmadas muestra que no considera que en todos los casos la decisión que toma un agente o la acción que efectúa involucre necesariamente creencias. Para sostener esto, parece necesario suponer que las pasiones calmadas también responden adecuadamente a cierta información sobre las circunstancias del agente. En otras palabras, a pesar de la afirmación en contra de Hume, las pasiones en general sí tendrían ciertas cualidades representativas y son éstas las que despiertan los deseos característicos de las pasiones, por ejemplo, los que mencioné antes como distintivos de las pasiones calmadas. El origen de estas representaciones es, según el pensamiento de Hume, desconocido (cf. T 1.1.2.1 SB 7), pero dado su modelo general de la mente, tendrían que ser alguna forma de percepción relativa a las impresiones sensibles. En consecuencia, me parece que puede aceptarse que tales representaciones no son necesariamente creencias y, por lo tanto, no son así objeto de la razón. Aun cuando este argumento pueda explicar el efecto de las pasiones calmadas sobre la voluntad es insuficiente para responder a Korsgaard, porque tal como ella lo ha planteado, la verdadera irracionalidad tiene que admitir la posibilidad de que un agente falle en responder apropiadamente a una creencia disponible.

De acuerdo con su crítica a Hume, dicha creencia disponible tendría que ser aquella que, bajo la forma de un juicio, acompaña a una pasión y puede referirse o a un objeto que da fundamento a una pasión como miedo, pena, alegría o desesperación, o bien puede referirse a los medios para alcanzar el fin de una cierta pasión que se ha puesto en acto (cf. R2, v. supra p. 80). Para Hume estos dos tipos de juicio influyen sobre la pasión sólo en la medida en que son erróneos, o bien porque el objeto de la pasión no existe o bien porque los medios para alcanzar el fin son insuficientes. Ahora bien, la objeción de Korsgaard no parece dirigirse correctamente al primer caso.

El tipo de creencia en este caso se refiere al objeto de una pasión, particularmente, al objeto del miedo, de la pena, de la alegría, etc. Considerando esto, no es extraño que la pasión «se someta a la razón sin oposición alguna» porque ni el miedo, la alegría o la desesperación pueden entenderse sin la relación adecuada con un cierto objeto. Si un agente experimenta miedo porque cree que está a punto de ser atacado por un perro es comprensible que su miedo se desvanezca inmediatamente si de pronto descubre que se trata de un perro amistoso y juguetón. Nos parecería una anomalía preocupante si no sucediera así. En otra variable del caso, la alegría de alguien

porque ha creído entender que recibirá la visita de una persona muy anhelada se disipa inmediatamente cuando comprende que ha entendido mal algo y que no recibirá dicha visita. No cabe esperar que la alegría continúe si desaparece el objeto que la despertó. En consecuencia, no cabe suponer que la verdadera irracionalidad como la entiende Korsgaard aplique a estos casos. Fallar en responder a estas creencias –en vez de someter las pasiones a ellas como dice Hume– equivaldría a mantener la tristeza o la alegría (o cualquier otra pasión) sin un objeto inteligible que les de sustento o posiblemente equivalga a un grave caso de auto-engaño, el cual tampoco es el tema que le interesa a Korsgaard.

El segundo caso se refiere a la creencia errónea que tiene un agente sobre cómo alcanzar el fin marcado por poner en acto una pasión. Supongamos que la pasión del agente es el miedo por lo que considera la amenaza inminente de ser atacado por un perro. Esta pasión se pone en acto haciendo algo para alejarse del objeto del miedo, de modo que el agente decide caminar hacia lo que considera un lugar seguro. Se equivoca en esta valoración y resulta que no es un lugar seguro. De modo que cabe esperar alguna corrección en su decisión. Hume en este punto presenta la célebre consideración sobre la incapacidad de la razón para guiar a la pasión en la dirección apropiada (cf. T 2.3.3.6 SB 416; v. arriba R3, p. 81). Pero si antes de volver a examinarla traemos a cuento la objeción de Korsgaard, de nuevo me parece que no aplica adecuadamente en este caso. Para empezar, no cabe esperar que el agente fallara en actuar de acuerdo con la mejor creencia disponible para él, a saber, que el lugar que ha elegido no es seguro. Si a pesar de tener esta creencia se empecina, digamos que por probar un punto sobre la irracionalidad, en mantenerse en ese lugar las consecuencias podrían ser terribles y con todo habría probado el punto en cuestión. En este caso, ilustrar la verdadera irracionalidad de Korsgaard no es una alternativa viable y por ello no aplica. Con todo, Korsgaard no se refiere seguramente a situaciones como ésta, sino a aquellas en las que ciertas garantías de seguridad para el agente salvadas, éste puede permitirse alguna irracionalidad. En efecto, ella además considera que una teoría de la motivación tiene que satisfacer el requisito internista, el cual consiste en: “Las afirmaciones de razón práctica, si realmente se nos presentan con razones para actuar, tienen que ser capaces de motivar a las personas racionales” (Korsgaard 1986, 318). Es decir que un agente racional, en tanto tal, encontraría en ciertas razones sobre qué hacer motivos para actuar de acuerdo con ellas. Esto no implica que «nada pueda interferir con la transmisión motivacional», porque, según nos dice Korsgaard, “La necesidad, o la obligatoriedad [*compellingness*], de las consideraciones racionales

yace en aquellas consideraciones mismas, no en nosotros: esto es, nosotros no estaremos necesariamente motivados por ellas” (Korsgaard 1986, 320). De modo que el agente puede, como hemos examinado ya, fallar en actuar de acuerdo con una razón apropiada para una situación dada. En esto consiste la verdadera irracionalidad, en que el agente podría fallar, no que necesariamente fallara. Ahora bien, el agente humeano que se equivoca de medio para escapar de una amenaza, corrige su decisión y actúa de acuerdo con una consideración que lo pone a salvo. De este modo satisface el fin de la pasión puesta en acto. Con todo, la razón no es, según Hume, la facultad que ha operado esta resolución práctica, sino una pasión calmada. Su explicación, me parece a mí, es más amplia e interesante de lo que parece cubrir el requisito internista de la racionalidad práctica de Korsgaard.

Para empezar, la explicación de Hume contiene, me parece, algunos elementos que es necesario presentar más claramente. La elección de términos usados por Hume no deja ver que en realidad en los casos en los que hay una creencia errónea sobre los medios para satisfacer una pasión, inmediatamente después de que «la pasión se somete a la razón», o bien en el momento mismo, surgen por lo menos dos creencias candidatas a dirigir el curso de acción. En el memorable ejemplo del *Tratado* estas creencias son los disyuntos de: «O se destruye el mundo o tendré un rasguño en el dedo». Si la situación se reconstruye agregando elementos para que la disyunción sea plausible, podría entenderse que ante el miedo a sufrir un daño, esto es, la expectativa de sufrir un rasguño, el agente considera otra alternativa, la expectativa de salvarse a sí mismo a costa de destruir el mundo. Ahora, el punto importante de Hume es que la razón no tiene forma de valorar la deseabilidad de cada elemento de la alternativa. Hume insiste en esto: “No es contrario a la razón que yo prefiera mi ruina total con tal de evitar el menor sufrimiento a un *indio* o a cualquier persona totalmente desconocida. Tampoco es contrario a la razón el preferir un bien pequeño, aunque lo reconozca menor, a otro mayor... (T 2.3.3.6 SB 416). La razón, por sí sola, no tiene recursos para decidir sobre la deseabilidad de los disyuntos de estos dilemas. Desde su punto de vista ambos disyuntos tienen el mismo valor –si bien Hume parece decir que incluso podría tener los valores invertidos: preferir la destrucción del mundo o un bien menor, etc. Así que si todo dependiera de la razón estaríamos en problemas. Es aquí donde entran las pasiones calmadas y «su deseo instintivo por la benevolencia, el resentimiento, el amor a la vida y la ternura para con los niños o su apetito general hacia el bien y la aversión contra el mal». En efecto, a las pasiones calmadas pertenece nuestro sentido moral y estético, según lo ha puesto Brand (cf. Brand, 24). De

modo que, aunque hay dos creencias en conflicto, éste no se resuelve por un procedimiento de razonamiento práctico sino por un procedimiento *deliberación* práctica. Esto es, según Hume no es por medio de la reflexión racional como los agentes deciden qué pensar o qué hacer en una situación dada, sino que por medio de diversos elementos mentales toman una decisión o actúan de alguna manera. Siguiendo el uso común al respecto, me referiré a este proceso como *deliberación práctica* (Blackburn 1998, Gibbard 1990, Gibbard 2003, Taylor 1985, Raz 1999, Raz 2011, Wiggins 1987).

La característica distintiva de la deliberación práctica es que la razón no es capaz de decidir sobre los fines de la acción y del pensamiento. Como hemos visto, en el caso de Hume, afirmar esto consiste en otra forma de su escepticismo sobre las capacidades de la razón. Sin embargo, ésta es una posición negativa respecto de las funciones de la razón y podemos inquirir, en consecuencia, por los rasgos positivos distintivos de la deliberación práctica. Al respecto, Hume propone un modelo persuasivo de la deliberación práctica que deja a las pasiones la tarea de establecer los fines de la acción y del pensamiento. Por ejemplo, ante un conflicto de creencias como el presentado arriba, la pasión se inclina por una de las dos creencias y resuelve así el conflicto. Estas pasiones pueden ser también las pasiones calmadas y aunque esta solución puede parecer *ad hoc* para evitar que sea la razón la responsable de decidir sobre algún fin su caracterización de tales pasiones es, me parece, plausible. En efecto, las pasiones calmadas consisten en ciertos principios e instintos que predisponen a los agentes a preservar su bien, a procurarse belleza y los objetos de su amor, incluyendo el amor propio. Además, podemos aceptar también que con esta teoría Hume es un precursor de la idea de que hay contenidos mentales inconscientes y que no siempre es irrelevante mencionarlos para explicar muchos de nuestros pensamientos y acciones.

A estos elementos, podemos sumar el papel que la imaginación, el hábito y la costumbre tienen en el sistema de Hume para explicar estados mentales. No veo ninguna restricción teórica para que estos elementos desempeñen también un papel en la deliberación práctica –en cambio, sí hay impedimentos para que la razón pueda jugar un papel distinto al de un mero instrumento de la pasión. Como hemos visto, estas determinaciones de la mente son las responsables de prácticamente todas las creencias y del paso de una pasión a otra, de modo que su función en la formación de pensamientos y su influencia sobre la voluntad parece innegable. Creo que una

reelaboración de estas ideas, más sencilla y tal vez más clara, la podemos encontrar en las ideas de Bernard Williams sobre la motivación (v. Williams 1993).

B. Williams defiende un «modelo» sobre la motivación que atribuye a Hume, aunque como su versión le parece más simple que la del propio Hume, la designa como *subhumano* (cf. Williams 1993, 132). En su presentación inicial, este modelo consiste en las siguientes afirmaciones:

- A tiene una razón para φ si A tiene algún deseo que se verá satisfecho si A hace φ .
- Algún deseo que A cree que se verá satisfecho si hace φ .

De acuerdo con Williams, ésta es la interpretación *internista* más simple del *hecho* de que un agente (A) tenga una razón para actuar –denotado por φ , que representa un verbo de acción (que Williams presenta con los enunciados “A tiene una razón para φ ” o “Existe una razón para que A φ ”). Según él, este hecho puede interpretarse de dos maneras, una internista y la otra externista. La primera sostiene que «el agente tiene una motivación interna que se verá satisfecha o favorecida por el hecho de que haga φ . De modo que el internismo es una especie de requisito que establece que tiene que haber una relación entre la acción y los objetivos del agente para que éste tenga una razón para φ . Por contraste, en la interpretación externista no existe este requisito y aun así el agente tendría una razón para φ . En otras palabras, el externismo aceptaría que un agente puede tener una razón para φ aun cuando φ no tenga ninguna relación con los objetivos del agente». También se pueden presentar estas interpretaciones examinando los enunciados mencionados arriba. Según la interpretación internista, cualquiera de los enunciados en cuestión será verdadero si y sólo si se satisface el requisito en cuestión, a saber, que haya una relación entre hacer φ y los objetivos del agente. Sin esto es falso. Por contraste, para la interpretación externista el enunciado «no será falso por la ausencia de una motivación adecuada», esto es, la ausencia del requisito en cuestión no hace necesariamente falso que ‘el agente tenga una razón para φ ’ (cf. Williams 1993, 131). Ahora podemos ver por qué Williams designa a su modelo de interpretación como sub-humano.

Según esta interpretación del internismo –Williams nos advierte que puede haber otras- el requisito en cuestión se presenta como un deseo, de modo que un agente tiene una razón para hacer φ sólo si hacerlo contribuye a la satisfacción de un deseo suyo. En último término, lo que motiva a un agente y le da razones para actuar es un deseo. En consecuencia, la fuente de la

motivación no es una creencia ni un principio ni una orden o una recomendación o un consejo, etc. Así, esta teoría de la motivación requiere *necesariamente* de un deseo del agente para que éste tenga una razón para actuar. (Smith ha identificado en esta idea uno de los dogmas más arraigados de la filosofía contemporánea, v. Smith 1994, cap. 4, p. 92). Con todo, si el internismo se limita a este modelo subhumeano deja mucho que desear. En efecto, el propósito de Williams es «desarrollar este modelo por adición y hacerlo más adecuado» (Williams 1993, 132). Me parece que las dificultades más apremiantes surgen de la necesidad de clarificar al deseo, porque no parece claro cuál podría ser su fuente o cómo se caracterizan exactamente ni cómo se relacionan con otros estados mentales que podrían ser relevantes para explicar a la motivación. Williams pronto reemplaza esta primera formulación centrada en el deseo por una más abarcante, en donde el requisito internista se relaciona no exactamente con un deseo del agente si no con algún miembro de lo que llama *conjunto motivacional subjetivo* (*S*). De este modo, el requisito internista puede expresarse del siguiente modo: un agente tiene una razón para φ si y sólo si φ se relaciona *apropiadamente* con algún elemento del *S* del agente. La cláusula *adecuadamente* permite, entre otras cosas, excluir «a los elementos de dicho conjunto basados en creencias falsas» (cf. Williams 1993, 132). Para hacer más claros los vínculos con Hume, la idea es que un agente no tiene una razón para φ si el elemento del *S* con el que se relaciona se refiere a un objeto inexistente o si φ no es un medio apropiado para satisfacer al elemento relativo del *S*. Por ejemplo, si una persona desea satisfacer su sed y cree que beber del vaso frente a ella es el medio adecuado para satisfacerla porque tiene agua tiene una razón para beberlo si y sólo si el vaso efectivamente tiene agua. Si en lugar de ello tiene, digamos, gasolina, entonces no tiene una razón para beberlo. Lo que la persona estima como el medio para satisfacer al elemento de *S* no le da una razón para actuar en este sentido, porque hay una creencia falsa de por medio. Con base en esto, Williams observa que su modelo reformado mediante *S* tiene dos consecuencias epistémicas importantes: a) Un agente puede creer erróneamente un enunciado sobre una razón interna respecto a sí mismo y b) Este agente puede no conocer un enunciado sobre una razón interna respecto de sí mismo (cf. Williams 1993, 133). La conclusión (b) nos hace presente la idea de que muchos estados mentales podrían no ser conocidos para el agente que los posee, por ejemplo porque están en el inconsciente, y aun así ser relevantes para explicar su acción o pensamiento. Quizá más importante, para la teoría de la deliberación de Hume, es que con estas conclusiones de Williams

podemos ahora expresar dos aspectos centrales de toda teoría de la motivación: el aspecto psicológico y el aspecto normativo de las razones para actuar.

Con el primero de éstos me refiero al hecho de que un agente muchas veces actúa sobre creencias o información falsa: Como *creía* que el líquido era agua, tomó del vaso. Con el aspecto normativo quiero hablar de que hay restricciones normativas sobre los motivos, en este caso, que no pueden estar basados en creencias o suposiciones falsas. Así, muchas veces sucederá que el aspecto normativo afecta adecuadamente al psicológico, precisamente en los términos en que lo ha puesto Hume (y que cuestiona Korsgaard): a la luz de información falsa o en cuanto descubrimos la falsedad de una creencia que nos parecía relevante para nuestra acción, el deseo (o cualquier otro elemento de *S*) cesan o se ajustan a la nueva información. Así, en cuanto el agente se percató de que hay gasolina en el vaso, su motivo para beber de él desaparece. Ahora bien, sólo por una cierta presión teórica diríamos que ésta es la respuesta *racional* a esta situación, pero en realidad el ajuste motivacional se hace, según Hume, de manera inmediata, sin mediación de ningún proceso reflexivo. Por su parte, Williams cuenta con un recurso para incluir cómo esta relación entre el aspecto normativo y el psicológico excluye eficazmente a la información o creencias falsas, esto es, para que el aspecto normativo influya adecuadamente sobre el psicológico: “En el razonamiento deliberativo pueden descubrirse los enunciados sobre las razones internas” (cf. Williams 1993, 134). Parte de la importancia de esta idea consiste en que hace lugar a procesos deliberativos que permiten a los agentes hacerse conscientes de sus razones para actuar o, desde un punto de vista en tercera persona, que permite examinar los méritos epistémicos de las razones para actuar y aceptarlas o rechazarlas. Además, la distinción entre razonamiento *deliberativo* y razonamiento *práctico* muestra dimensiones diferentes de los procesos mentales detrás de una razón para actuar.

Esta distinción descansa en la observación de que no todos los procesos por los cuales se «descubre que un cierto medio de acción es el medio causal para un fin no es por sí mismo un razonamiento práctico». En efecto, podemos descubrir por suerte o por un proceso de libre asociación o en medio de una ensoñación altamente imaginativa cuál podría ser una forma de alcanzar un cierto fin que nos interesa. Por contraste, de acuerdo con Williams el razonamiento práctico se limitaría a aquel proceso inferencial de la forma general:

...que lleva a la conclusión de que uno tiene razones para φ porque φ podría ser la manera más conveniente, económica, placentera, etc., de satisfacer algún elemento de *S*, y esto, por supuesto,

está controlado por otros elementos de *S*, aunque no necesariamente de manera clara o determinada (Williams 1993, 134).

Así, el razonamiento práctico está regulado, de alguna manera, por la racionalidad relativa a la satisfacción de la eficiencia, la economía, el placer, etc., de un objetivo del agente. De este modo, es un razonamiento instrumental por medio del cual el agente descubre el vínculo causal que sigue alguna de las pautas de la racionalidad mencionadas. Pero los procesos deliberativos son más amplios y no se restringen a pautas de racionalidad, sino que suelen incluir elementos diversos. Williams describe así tales procesos:

Pensar cómo se puede combinar la satisfacción de los elementos de *S*, por organización temporal, por ejemplo; cuando existe algún conflicto irresoluble entre los elementos de *S*, considerar a cuál se le da más peso (lo que no implica, y esto es importante, que haya un bien del cual proporcionan cantidades distintas); o nuevamente, encontrar soluciones constitutivas, como cuando se decide cómo se lograría pasar una tarde entretenida, suponiendo que uno quiere entretenerse (Williams 1993, 134).

En conjunto, estas ideas muestran que los procesos deliberativos son dinámicos y *fluidos*, en el sentido en que como no están necesariamente sujetos a pautas de racionalidad pueden relacionar libremente diversos elementos de *S*. Esto explica cómo el agente descubre razones para actuar que no sabía que tenía, en atención a deseos, intereses, antojos, etc., que pueden presentarse a su atención con mayor o menor intensidad a través de estos procesos deliberativos. La combinación dinámica e *imaginativa* de los diversos elementos de *S* hace posible esto (cf. Williams 1993, 135).⁶⁰ Ahora bien, hemos visto que el modelo sub-humano de la interpretación internista presupone que el elemento necesario para explicar a la motivación es un deseo. Williams señala que aun cuando todos los elementos de *S* pueden entenderse *formalmente* como deseos, precisa que: “*S* puede incluir cosas como disposiciones de evaluación, patrones de reacción emotiva, lealtades personales y diversos proyectos, como podría llamárseles en forma abstracta, que incluyen compromisos del agente” (Williams 1993, 135). Con esta ampliación y especificación de los elementos de *S* podemos entender cómo el proceso de deliberación del internismo que propone Williams no es susceptible de ser caracterizado racionalmente.

En los términos en que Williams ha trazado la distinción entre razonamiento práctico y razonamiento deliberativo, el primero puede caracterizarse como un proceso inferencial mediante

⁶⁰ El contexto de esta referencia alude a la complejidad del proceso deliberativo y cómo a través de él un agente puede descubrir razones para actuar que no sabía que tenía. En este contexto, la imaginación desempeña un papel al “crear nuevas posibilidades y deseos”. Williams sostiene con base en este que «éstas son posibilidades importantes tanto para la política como para la acción individual».

el cual un agente determina el mejor curso de acción para realizar algún proyecto propio de acuerdo con una pauta racional relativa a dicho proyecto –satisfacer un deseo, honrar un principio, actuar de acuerdo con una cierta idea, etc. En otras palabras, este proceso inferencial sólo sirve instrumentalmente a ciertos fines.

La cuestión de si la razón, por medio del razonamiento práctico, puede determinar el fin de la acción de un sujeto o el objetivo de ciertos proyectos, en vez de un deseo, me parece que puede presentarse bajo el examen que Williams hace del siguiente problema, relativo a los bienes públicos y de quienes se benefician gratuitamente (*free riders*) de ellos (el interés de Williams, en esta cuestión, parece ser, sobre todo, examinar si es irracional actuar bajo motivaciones egoístas y su respuesta es que no siempre lo es). La situación en cuestión es que cada persona tiene razones egoístas para querer cierto bien disponible y al mismo tiempo razones egoístas para no tomar parte en el aprovisionamiento de dicho bien (cf. Williams 1993, 143). Me parece que este problema es equivalente, bajo ciertas modificaciones, al problema de si podemos aceptar que la razón puede determinar el fin de nuestras acciones. Para empezar, el problema se refiere a una situación en la que un agente dado tiene motivos o razones egoístas para hacer algo o no hacerlo, a saber, aprovechar un bien público y no aprovecharlo. La resolución de este dilema es el objetivo de la razón. Suponemos que mediante algún procedimiento racional, la razón derivara la conclusión que soluciona el dilema y que de este modo determina un fin que el agente adopta. En otras palabras, la razón puede establecer una razón para actuar en relación con un dilema del tipo o hago φ o no hago φ , pero no ambas e independientemente de cualquier otro interés que pueda tener, por ejemplo, desear hacer φ . Ahora bien, Williams no se propone «discutir este problema», sino examinar cómo se podría aclarar su posición sobre las razones internas considerando algunas preguntas relativas al problema desde la perspectiva de un defensor de «la racionalidad de las razones internas», es decir, de un defensor del internismo. Yo no consideraré directamente las preguntas de Williams, pero presentaré sus consideraciones de manera argumental. Mi objetivo será responder, con base en sus consideraciones, si puede defenderse que la razón puede determinar el fin de la acción.

De acuerdo con Williams, (1) podemos definir nociones de racionalidad que no sean puramente egoístas y también (2) podemos definir nociones de racionalidad que no sean puramente medios-fin, pero (3) no podemos definir ninguna noción de racionalidad en la que ésta no se refiera a las motivaciones existentes de un agente dado. Estas primeras observaciones de

Williams parecen equivaler a que podemos entender a la racionalidad dirigida, por ejemplos, al interés de otras personas o al interés común de un conjunto de personas asociadas entre sí de alguna manera. Por ejemplo, mediante una consideración sobre qué podría beneficiar más a tal persona en una situación dada o qué acciones de un agente dado podrían aportar un beneficio a dicho conjunto de personas, aun cuando esas acciones puedan afectar a sus propios intereses. De esto podemos derivar una razón a favor de su segunda observación: si un agente puede hacer algún *cálculo* sobre cómo beneficiar a otra persona o a un conjunto de personas, restando del mismo sus propios intereses, entonces este cálculo es un caso de razonamiento sobre los fines y no meramente sobre los medios. Con todo, la definición (3) requiere que este cálculo se refiera a las motivaciones previamente existentes del agente, de modo que cualquier acepción que podamos ofrecer de ‘no egoísmo’ o de ‘restar los propios intereses’ tendrá que satisfacer este requerimiento. Esta observación sugiere que cualquiera que sea el resultado del cálculo racional, éste, en primer lugar, está motivado previamente por algún elemento del *S* del agente –por ejemplo, *querer* resolver racionalmente el dilema en cuestión. De este modo, sin embargo, parece que (3) contradice a (2), porque (2) decía que puede definirse una noción de racionalidad medios fines, pero si el proceso racional que nos interesa es despertado por un elemento previo del *S* del agente, entonces en realidad *sirve* de medio a dicho elemento. Por medio del cálculo racional se satisface ese elemento del *S*. Esta aparente contradicción entre (2) y (3) nos requiere explorar con más detalle y recurrir a otros elementos. Dada que el argumento general de Williams favorece la aceptación de la interpretación internista de las razones para actuar, (3) tiene que entenderse como una implicación de tal argumento. De este modo, las consideraciones sobre las nociones de racionalidad no puramente egoístas, (1), tienen que entenderse del lado del externismo. Esto es, el externista tendría que ofrecer alguna prueba a favor de lo que sería tener un razonamiento sobre un fin propio que no esté relacionado con nuestros intereses pre-existente. Esta misma consideración tendría que valer para la racionalidad que no sea puramente medios-fines. Es el externista quien tendría que defender una racionalidad que funcione correctamente sobre el fin de la acción y que tal proyecto se pueda entender al mismo tiempo sin recurrir a los intereses antecedentes del agente, de otro modo, como argumenté arriba, se trataría nuevamente de la racionalidad instrumental. Williams no estima que pueda entenderse adecuadamente la idea de tener una razón externa:

El tipo de consideraciones que ofrecemos aquí me sugieren enfáticamente que los enunciados sobre razones externas, cuando están definitivamente aislados como tales, son falsos, o incoherentes, o en realidad expresan erróneamente otra cosa. De hecho, es más difícil aislarlos en el habla de las

personas de lo que se sugirió en la introducción de este capítulo [“Razones internas y externas”]. Quienes utilizan tales palabras a menudo parecen aceptar, más bien, una afirmación optimista de razones internas, pero a menudo el enunciado se presenta en realidad definitivamente fuera del *S* del agente y lo que él podría derivar de ella mediante deliberación racional; y entonces, sugiero, es sumamente incierto lo que se quiere decir. Algunas veces es tan sólo que las cosas serían mejores si el agente actuara así (cf. Williams 1993, 142).

En consecuencia, aunque desde el punto de vista el externista podría pensarse que es plausible un razonamiento sobre un posible motivo o razón que no esté relacionado con los interés previos de la persona –en este sentido se puede definir a la racionalidad no egoísta, según acepta Williams con (1)– en realidad, según Williams, parece muy difícil caracterizar el espacio psicológico de este proceso, es decir, no parece que pueda entenderse una reflexión pertinente para un cierto agente sobre un cierto interés suyo que no forma parte de sus intereses o que no sea relativo a su *S*. Con la última frase que cito arriba, Williams parece sugerir que el externista podría estar así moralizando para promover sus preferencias sobre la conducta de otros agentes. Con todo, esta última observación pertenece al contexto más general del argumento de Williams sobre la implausibilidad de interpretar correctamente a las razones externas. En el contexto de la contradicción que estoy considerando, la cuestión es que un agente podría tener buenas razones para actuar no-egoístamente respecto de un bien público. Williams discute con esto una forma muy abstracta del problema del *gorrón* o *free rider*.

Este problema se refiere a las condiciones en que un grupo de agentes asociados entre sí puede acceder a un cierto recurso muy deseable pero finito (si pudiera disponerse sin límite alguno del recurso no habría problema). En estas condiciones lo racional parece ser que todos acuerden restringir su acceso al recurso o compensar de algún modo su acceso, con el fin o bien de prolongar lo más posible su uso o bien ofrecer alguna participación para producir una alternativa viable. Así, aunque todos tienen un interés egoísta en el recurso, porque conviene a intereses particulares, al mismo tiempo es racional no actuar meramente por motivos egoístas, sino en favor de los intereses de todo el grupo. En este punto puede surgir un gorrón que no restringe sus intereses egoístas, ya sea porque no limita su acceso al recurso o porque no participa con la compensación (tal vez no paga cuotas o impuestos). Si el gorrón fuera un caso aislado, es posible que la comunidad de agentes pueda soportar su conducta no solidaria, al menos por un tiempo. Pero si muchos más agentes se convierten en gorriones entonces en algún momento el recurso se agotará y nadie podrá ya disfrutar de él ni tener una alternativa viable. Es este condicional el que muestra que es irracional en este contexto actuar meramente por motivos egoístas. Al mismo tiempo, Williams

observa que incluso en este punto no siempre es irracional actuar por motivos egoístas. Particularmente, «en aquellos casos en que su propia participación en las prácticas solidarias es decisivo para que éstas sean suficientemente eficaces o al menos que al agente así lo crea» (cf. Williams 1993, 143, la respuesta a la pregunta 6). Esto es, si incluso con su participación no se alcanza el número crítico de personas para revertir la influencia de los gorriones, entonces el agente no vería ningún motivo para actuar cooperativamente; por el contrario, si al hacerlo se alcanza ese número crítico, entonces actuar cooperativamente es, ahora, parte de su interés y tiene un motivo egoísta para hacerlo, digamos, seguir disfrutando del recurso en cuestión. Creo que lo importante de estas observaciones es que aún en este contexto no parece plausible entender que un razonamiento no egoísta pueda prescindir de elementos motivacionales antecedentes en la psicología de los agentes.

La última conclusión es muy significativa al respecto: un agente podría adoptar una conducta solidaria, en una situación dada, porque hacerlo sirve a sus intereses. Esto es la conducta racional para este agente y para el grupo, pero no parece que la decisión importante se pueda alcanzar al margen de ninguna consideración de los antecedentes, situación del agente, etc., que pueden incluir, como en este caso, referencias a los intereses de otros agentes particulares o a los intereses colectivos de una comunidad de agentes. Si esto es correcto, entonces ésta parece ser la forma de entender la relación que (1) y (2) tienen con (3): habría una contradicción entre ellas sólo si suponemos que la deliberación racional no egoísta se entiende bajo una interpretación externista de las razones para actuar. Bajo una interpretación internista no hay tal contradicción porque cualquier caracterización que hagamos de las razones no egoístas, incluso con la cláusula “no puramente egoístas”, implica alguna referencia relevante al *S* del agente, esto es, siempre la entenderíamos en términos de intereses preexistentes del agente y con suficiente fuerza motivacional para despertar el proceso de razonamiento no-egoísta. Con base en estos elementos, me parece, puede ofrecerse una respuesta a las objeciones de Korsgaard a Hume y al propio Williams.

En su reconstrucción del argumento de Williams, Korsgaard le atribuye una posición escéptica parcial que se refiere a que un razonamiento práctico tiene que ser capaz de motivar a un agente, mientras que *no* es escéptico respecto de que la razón sólo procede de medios a fines (cf. Korsgaard 1996, 326). En la interpretación de Korsgaard, esto es una importante diferencia con Hume, porque según ella él es escéptico respecto de ambas capacidades de la razón.

Korsgaard reconoce que Williams sostiene que la razón es del tipo medios/fines por consideraciones similares a las que he presentado arriba:

Las razones externas... existen independientemente de lo que está en el conjunto motivacional subjetivo de uno. En este caso, Williams señala que tiene que haber un proceso racional, que no nace del conjunto motivacional subjetivo y por lo tanto que no es relativo a él, por el cual tú puedes reconocer que algo es una razón y al mismo tiempo estar motivado por ella. La razón tiene que ser capaz de producir un motivo completamente nuevo... (Korsgaard 1996, 326).

Como para un agente dado sólo le es posible razonar a partir de lo que ya está en su *S*, entonces no puede producir un motivo completamente nuevo, esto es, no puede proponerse actuar bajo lo que puede identificarse como una razón externa. Para Korsgaard esto equivale a una forma de escepticismo sobre la razón práctica que descansa en el *contenido* de las razones. Según ella, hay dos formas de escepticismo sobre la razón práctica. El primero, el escepticismo de contenido, consiste sostener que ninguna consideración racional tiene el *contenido* adecuado para ofrecer «guía substantiva para la elección y la acción». El segundo se refiere al alcance de la razón como motivo para la elección y la acción, esto es, que la razón es incapaz de motivar sobre los fines de nuestros pensamientos y acciones (cf. Korsgaard 1996, 311). Así, mientras que Hume sería, según ella, un escéptico de ambos tipos, Williams sólo lo es del segundo, porque para él, «la clase de ítems que pueden encontrarse en el *S* no limitan, sino que más bien dependen de qué tipos de razonamientos son posibles» (cf. Korsgaard 1996, 328). De este modo, para Korsgaard tanto Williams como Hume sostendrían un escepticismo sobre los fines que depende de su escepticismo de contenido (cf. Korsgaard 1996, 329). En el caso de Williams, no encontraremos en el *S* del agente ningún contenido que permita el tipo de plataforma psicológica para aceptar una razón externa, porque la propia razón requiere de elementos antecedentes, relativos al propio *S*. La respuesta de Korsgaard a esta posición es la siguiente, en la que también incluye a Hume:

El argumento de Williams no muestra que si hubiera principios incondicionales de la razón aplicados a la acción no podríamos estar motivados por ellos. Él sólo cree que no hay ninguno. Pero el argumento de Williams, como el de Hume, da la impresión de ir en la dirección contraria: parece como si el punto motivacional –el requerimiento internista– supuestamente tiene alguna fuerza para limitar qué cuenta cómo un principio de razón práctica. Mientras que de hecho, la fuente real del escepticismo es una duda acerca de la existencia de principios de acción cuyo contenido les muestran estar justificados últimamente (cf. Korsgaard 1996, 329).

Korsgaard parece estar pidiendo a Williams que una razón externa a un agente –un principio incondicionado de la razón, como un imperativo categórico kantiano– pueda motivarlo. A mí me parece que nada en lo que ha dicho Williams, o Hume, niegue la existencia de tales razones, sólo que éstas, por sí solas, no tienen ninguna fuerza motivadora, a menos, en el caso de ambos, que

tales razones sirvan de medio para un fin previamente determinado por un deseo o por una pasión. Según la posición de ambos, no puede haber ninguna razón que tenga fuerza motivadora totalmente desconectada de, formalmente, un deseo del agente. Tal como Korsgaard entiende el internismo, esta posición es inaceptable. Según ella, el internismo sólo requiere que las afirmaciones de la razón práctica, si realmente nos presentan razones para actuar, tienen que ser capaces de motivar a personas racionales (cf. Korsgaard 1996, 319). Entendido así, el requisito internista no requiere de deseos, sino sólo que los agentes racionales, en tanto racionales, respondan adecuadamente a una razón disponible para ellos. Pero esto no nos dice nada sobre el agente particular ni sobre la situación en la que esperamos que actúe. Este agente racional, susceptible de ser motivado por una razón práctica disponible, parece que podría suspender cualquier interés, consideración particular, deseo que pudiera tener, etc., y actuar sólo bajo la guía de tal razón práctica. Creo que la dificultad que ha indicado Williams, para entender tales razones, es aquí realmente insuperable (v. Williams 1993, 142, supra 105 y ss). Williams, si en esto puede ser un vocero adecuado del modelo humeano de motivación, nos estaría recordando que los agentes humanos no son meramente racionales y que sólo podemos entender sus capacidades racionales, de un modo instrumental, dentro del entramado de otros elementos de la psicología de los agentes.

3.5. El argumento de Smith para defender una interpretación humeana de la motivación: «el observador ideal»

Hasta ahora, con las consideraciones de Williams, podemos entender cómo la teoría de la motivación de Hume satisface el requisito de un agente mínimo pero eficiente para deliberar sobre sus pensamientos y acciones –es mínimo porque la razón sólo le ayuda instrumentalmente, no sobre los fines de la acción. De este modo parece responderse el cuestionamiento de Korsgaard dirigido particularmente a Hume sobre la capacidad de la razón para servir instrumentalmente. También vimos cómo puede responderse a la idea de que este agente no cumple adecuadamente con el requisito de la verdadera irracionalidad, esto es, actuar incluso en contra de una razón verdadera. Hay todavía una cuestión pendiente, relativa a los fines de la acción: ¿puede disponer este agente mínimo de otros recursos evaluativos para sopesar los fines de su acción? Para responder esta cuestión desarrollaré ahora algunas ideas de Michael Smith, quien ofrece una respuesta a esta pregunta a partir de las dificultades comunes que enfrenta la teoría de la

motivación de Hume. En la última sección de este capítulo, dedicada al tratamiento que hace Hume de los juicios morales, ajustaré su teoría a los términos e interés propios de Hume.

M. Smith en “The Humean Theory of Motivation” (1987) se propone ofrecer un argumento a favor de la teoría *humeana* de la motivación –a pesar de que reconoce que se ha convertido en un *dogma* de la psicología filosófica y que ha encontrado numerosos e importantes detractores-, la cual consiste en la siguiente afirmación: la motivación tiene su fuente en un deseo relevante y en una creencia medios fines (cf. Smith 1987, 36; Smith 1994, 1994 y ss.). Así caracterizada, en efecto, el modelo explicativo de la motivación que hemos visto en Williams es un ejemplo de esta teoría –y también, hasta aquí, la propia teoría que leemos en el *corpus* de la obra de Hume es otro ejemplo de teoría *humeana* de la motivación. Posteriormente, en su importante *The Moral Problem* (1994), acepta que la distinción básica de la psicología *humeana* entre dos tipos de estados psicológicos es correcta, esto es, entre creencias y deseos (cf. Smith 1994, 7 y ss; v. también 93). De modo que hay una brecha categorial entre unas y otros que también separa a las creencias de los motivos –pues éstos responden a los deseos.⁶¹ De este modo, el problema moral se puede explicar cómo la separación que hay entre las razones morales para actuar y los motivos de los agentes, los cuales no necesariamente se refieren a asunto morales. Así, Smith acepta los elementos básicos de la psicología filosófica, que en términos más abstractos podemos ahora señalar como la separación categorial que hay entre estados cognitivos, las creencias, y los estados no-cognitivos, las pasiones y los deseos. Sin embargo, en *The Moral problem* Smith rechaza que este pecado de origen determine por completo a la teoría de la motivación: “... argumentaré que no podemos y no debemos aceptar la afirmación de Hume de que los deseos están por sí mismos más allá de toda crítica racional” (Smith 1994, 14). Esto es, Smith pretende defender que la razón puede influir sobre los deseos de manera que se produzcan motivos susceptibles de evaluación racional. Es justamente por este propósito que recurro a él, porque nos ofrece una vía para responder a la pregunta por la deseabilidad de los fines. Sin embargo, aceptaré con reservas su idea de que los deseos reciben bien la crítica racional, como diré a lo largo de esta sección.

Ahora bien, este proyecto desemboca en una versión del internismo que, según Smith, es diferente al de Nagel, McDowell y Platts –pero también es diferente al internismo de Korsgaard

⁶¹ V. supra nota 52 en p. 82. Platts señala precisamente esta brecha, pero la rechaza argumentando que las razones pueden despertar a los deseos, esto es, superar la brecha en cuestión.

presentado antes. Estos internistas, en la lectura de Smith, están de acuerdo en la siguiente tesis anti-humeana:

Si un agente juzga que es correcto para ella hacer ϕ en las circunstancias C, entonces ella está motivada a hacer ϕ en C (v. Smith 1994, 61).

Es decir, un juicio motiva sin más (Smith se refiere a juicios morales), particularmente, sin la presencia de un deseo. Smith califica a esta posición como incorrecta porque “nos compromete a negar, por ejemplo, que la debilidad de la voluntad y similares pueden derrotar la motivación de un agente mientras dejan intacta su apreciación de las razones morales” (v. Smith 1994, 61). Como hemos visto con Korsgaard, Smith pretende que casos como éste sean casos de genuina irracionalidad, de modo que defiende la siguiente versión del internismo, afín según él al pensamiento humeano:

Si un agente juzga que es correcto para ella hacer ϕ en las circunstancias C, entonces o ella está motivada a hacer ϕ en C o ella es prácticamente irracional (v. Smith 1994, 61).

Así, para Smith los casos de debilidad de la voluntad y «similares» son casos de irracionalidad, esto es, casos en los que un agente compromete su razón al no actuar de acuerdo con lo que ha juzgado como racionalmente correcto para ella. Ahora, si bien esta afirmación parece correcta, necesitamos considerar cuál es el argumento a favor de ella, dado que Smith cree que es consistente con la distinción humeana entre creencias y deseos. Según hemos visto, la posición de Hume no garantiza tal conexión, precisamente por la brecha mencionada entre los estados psicológicos. En otras palabras, como lo ha demandado Miller, Smith tiene que ofrecer una teoría substancial de la razón práctica que muestre la conexión que hay entre el juicio y la motivación. Según Miller, sin una teoría substantiva de la razón práctica la tesis de Smith es trivial y sin ésta su tesis más bien ofrece un argumento de peso a favor de la alternativa, es decir, la posición externista (cf. Miller 2003, 220).

El argumento de Smith en *The Moral Problem* corre a lo largo de estas estrategias generales: por un lado, un estudio de las razones y motivos para actuar, que muestra que para explicar correctamente a la deliberación práctica necesitamos revisar qué es un motivo para actuar. Smith distingue entre dos sentidos trenzados en las razones: el sentido normativo y el sentido explicativo (cf. Smith 1994, §4.2). El primero se refiere a la racionalidad o deseabilidad de una razón disponible para un agente y el segundo a la relevancia psicológica de una razón dada para explicar que un agente actuara de una cierta manera. Ahora bien, estos dos sentidos no

necesariamente se conectan, pero pueden hacerlo y cuando lo hacen los agentes se involucran en una revisión racional de sus motivos, estos es, una evaluación de la deseabilidad de sus fines. A mí me parece que Smith tiene razón en aceptar que una gran «variedad de razones de razones normativas», pero no parece necesario que se deriven de una revisión racional. Con todo, esto nos permite hablar de la segunda estrategia general, la cual se refiere a la teoría del observador ideal. Según él, las razones normativas requieren necesariamente de un espacio de deliberación plenamente racional. Ese espacio de deliberación es equivalente al punto de vista del observador ideal, si bien Smith no pone su teoría en estos términos. Sin embargo, creo que su teoría ajusta bien dentro de las líneas generales de esta teoría.

La tesis central de la teoría del observador ideal consiste en la afirmación de que un agente delibera sobre qué hacer en una situación dada basado en los sentimientos hipotéticos de un observador ideal, el cual reúne por lo menos las cualidades de estar completamente informado, se enteramente objetivo y no auto-interesado (cf. Radcliffe 1994, 37).⁶² De modo que creo que entre los teóricos que defienden este punto de vista podríamos incluir a Smith por su estrategia de la perspectiva deliberativa. Smith pretende defender una concepción humeana de la motivación o racionalidad práctica, preocupado por responder la aparente falta de racionalidad de la propuesta de Hume (cf. Smith 1994, 137).⁶³ La raíz del problema, como hemos visto, yace en la distinción de Hume entre estados-creencia y estados-deseo. Smith acepta que cada uno estos estados mentales es un «estado solo, unitario» y señala que en la distinción entre ambos estados reside “el valor en efectivo [*cash value*] de la doctrina humeana de que creencias y deseos son existencias distintas” (Smith 1994, 119). En otras palabras, Smith acepta que esta distinción es correcta y que es fundamental para explicar a la motivación humana (cf. también Smith 1987, §1). De otra manera, ésta es la distinción entre estados psicológicos, creencias y deseos, que presenté antes. El problema, dado el interés de Smith por establecer una posición humeana sólida de la racionalidad práctica, consiste en establecer satisfactoriamente la conexión entre ambas existencias distintas. En otras palabras, se trata de explicar cómo, a partir de la distinción dada entre estados

⁶² Más adelante, en la sección 4.3, explico que para Hume el *observador ideal* está plenamente informado en relación con el punto de vista de una cierta autoridad o sabio. El punto de vista del observador ideal reúne la autoridad de la experiencia y de dicha autoridad o sabio, y se constituye como un cierto estándar normativo.

⁶³ En varios otros textos M. Smith (1994, 2004, 2009) ha mostrado la plausibilidad del proyecto de contar con una noción plausible de racionalidad práctica humeana. Me concentraré exclusivamente en Smith 1994, por la afinidad con Hume entre la perspectiva deliberativa de Smith y la interpretación del observador ideal para el pensamiento de Hume.

psicológicos, podemos explicar que la deliberación de los agentes responda adecuadamente a las consideraciones racionales, por ejemplo, consideraciones relativas a la deseabilidad de un cierto fin que se proponen alcanzar. Smith considera que se puede defender que existe una conexión conceptual entre ambos estados psicológicos, a pesar de la distinción humeana inicial: “Por lo tanto todo se vuelve sobre qué, exactamente, es la naturaleza de la conexión causal entre creer que algo es valioso y desear que lo sea” (Smith 1994, 148). Al parecer, tal como lo ve Smith este problema está en el centro de la moral. Un poco antes de esta idea, observa señalado que:

Lo que está en cuestión es así la coherencia de la idea de en qué medida la deliberación es práctica sobre la base de nuestros valores. Como es quizá evidente, esto es justamente el problema moral de nuevo, reescrito para hacer explícito que lo que está en cuestión es la conexión entre la razón y la motivación. Y aquí, correspondientemente, encontramos el corazón del debate (Smith 1994, 137).

Este problema moral central, cuya solución ocupa precisamente a Smith, se puede expresar también como una tensión entre dos rasgos salientes de los juicios morales. Por un lado, los juicios morales presuponen objetividad (p.e., sobre qué es lo correcto) y, por otro lado, presuponen un sentido práctico (p.e. constituyen razones para actuar). Además, a la vista de la distinción de Hume entre estados-creencia y estados-deseo, cada uno de estos rasgos parece, según Smith, corresponderse con uno de tales estados. Así, el rasgo de la objetividad se corresponde con el estado-creencia. Suponemos que hay hechos morales y que nuestros juicios morales se refieren a ellos. El rasgo práctico se corresponde con el estado-deseo, de manera que un agente estaría motivado a hacer aquello que valora positivamente, *i.e.* un agente desea hacer lo que valora positivamente. La tensión que da origen al problema moral surge de que no es obvio cómo se relacionan el rasgo objetivo y el rasgo práctico de los juicios morales. Por ejemplo, no es obvio que lo que juzgamos como correcto constituya al mismo tiempo un motivo para actuar. No es incoherente ni manifiestamente absurdo que un agente *juzgue* que es correcto donar dinero para los niños de la calle y que, al mismo tiempo, no tenga el deseo o la motivación de hacerlo. El agente podría decir, ‘sé que es correcto donar dinero para los niños de la calle, pero no estoy convencido de tener una razón para hacerlo’ (cf. Smith 1994, 4-12). De manera que el problema moral refleja esta tensión entre la objetividad de los juicios morales y la psicología de la motivación de los agentes. No parece necesario que estos elementos se alienen alguna vez y, sin embargo, la moralidad parece requerir que lo hagan. Sucintamente, este problema puede expresarse mediante las siguientes proposiciones, aparentemente inconsistentes entre sí:

1. Los juicios morales de la forma ‘Es correcto que yo ϕ ’ expresan la creencia de un sujeto acerca de una cuestión de hecho objetiva, un hecho respecto de lo que es correcto hacer para ella.
2. Si alguien juzga que es correcto que ella ϕ , entonces *ceteris paribus*, ella está motivada para ϕ .
3. Un agente está motivado a actuar de cierto modo sólo en el caso en que ella tenga un deseo apropiado y una creencia medios-fines, donde la creencia y el deseo son, en términos de Hume, existencias distintas (Smith 1994, 12).

Smith explica que la inconsistencia surge porque (1) expresa mediante una creencia un juicio moral, la cual, de acuerdo con (2) está necesariamente conectada de algún modo con la motivación, con tener un deseo, según (3). “Así que (1), (2) y (3) implican juntas que hay alguna conexión necesaria entre diferentes existencias: la creencia moral y el deseo. Pero (3) nos dice que no hay tal conexión” (Smith 1994, 12). En efecto, (3) parece insoslayable e implica que, por ser deseos y creencias existencias separadas, no hay ninguna conexión necesaria entre la creencia moral y el deseo. Se hace presente así la inconsistencia mencionada entre estas proposiciones. Parece fácil, ahora, apreciar por qué Smith afirma que este problema también se expresa, desde el horizonte del *Tratado*, como la necesidad de explicar “la conexión entre la razón y la motivación” (cf. Smith 1994, 137). En esto consiste el problema moral, según la perspectiva de Smith, ¿cuál es su solución?

Como he dicho, Smith defiende que hay una conexión conceptual entre razón y motivos, o entre las creencias morales y el deseo (cf. Smith 1994, 148). A riesgo de simplificar demasiado su argumento –que en realidad recorre buena parte de su compleja respuesta al problema moral enunciado- me concentraré sólo en los aspectos que tienen que ver con la noción de razón para actuar.

Según mi reconstrucción del argumento, la estrategia de Smith defiende, en primer lugar, la idea de que una razón para actuar tiene tanto un sentido motivacional como un sentido normativo. Smith propone que la caracterización más fuerte de una razón para actuar es:

- P1 R en t constituye una razón motivadora de un agente A para que ϕ ssi existe algún ψ tal que R en t consiste en un deseo de A apropiadamente relacionado con ψ y una creencia tal que si ella ϕ ella podría ψ (Smith 1994, 92).

Es decir que, en un momento dado (t), un agente tiene una razón para actuar (R) si se encuentra en un estado mental (ψ) tal que la razón para actuar reúne apropiadamente tanto un deseo como una creencia que satisfacen dicho estado mental (ψ). Pero es claro que ‘apropiadamente’ solo

quiere decir que un cierto deseo y una cierta creencia tienen que constituir juntos una razón para actuar. A pesar de que aceptar P1 –o la idea de que una razón para actuar reúne apropiadamente un cierto deseo con una cierta creencia- constituya hoy en día una suerte de dogma de la psicología filosófica de la motivación, tanto para humanos como para anti-humanos, es necesario demandar una explicación sobre qué queremos decir con ‘apropiadamente relacionados’ (cf. Smith 1994, p. 93).⁶⁴ En atención a esta cuestión es que Smith introduce la distinción entre ‘razón motivadora’ y ‘razón normativa’, los dos sentidos presentes en una razón para actuar.

El sentido de la razón motivadora reside en su dimensión explicativa. Una razón motivadora *explica que* un agente haga ϕ o esté dispuesta a ϕ , *otras cosas iguales*. Así, mi razón para beber de este vaso de agua *explica que* beba del vaso o esté dispuesto a beber de él, a menos que algo cambie en la situación –me dicen que el vaso es de mi vecino, o que el agua no está limpia, o veo una mancha en el borde, etc., factores que me disuaden de beber del vaso.⁶⁵ La razón motivadora, dice Smith, es así *psicológicamente real*, esto es, constituye un estado psicológico genuino «que desempeñan un cierto rol explicativo en la producción de la acción». Por su parte, el sentido normativo de la razón para actuar consiste en alguna *verdad* de la forma ‘Es deseable o se requiere que yo ϕ ’. Así, tengo una razón para beber de este vaso, en su sentido normativo, si es verdad que es deseable que beba del mismo. Digamos, porque se me antoja tomar agua y en efecto el agua del vaso está limpia, disponible para mí, etc. Estas condiciones establecen que sea verdad que tenga una razón para beber del vaso y satisfacer así mi antojo. La distinción entre ambos sentidos coloca a cada tipo de razón en una *categoría* distinta: Mientras que el sentido motivador es un estado mental, un motivo psicológico, el sentido normativo se expresa mediante proposiciones de la forma “Es deseable o se requiere que A haga ϕ ” (cf. Smith 1996, 96). Así que haciendo *zoom* sobre los rasgos de una razón para actuar, el problema que estamos considerando aparece ahora como la pregunta acerca de cómo se conectan estos sentidos. No hay, nuevamente, ninguna imposibilidad lógica de considerarlos como separados. Yo podría tener una razón para beber agua del vaso y desconocer o no importarme que no sea agua limpia o, incluso, desconocer por completo que no contiene agua. Supongamos en efecto que no hay agua en el vaso en cuestión. Desde luego, si deseo tomar *agua*, no tendría razón alguna para beber del vaso. Con todo, si aún

⁶⁴ Desde luego, la respuesta a esta pregunta, *per se*, distingue a los humanos de los anti-humanos. En este contexto, Smith ubica entre estos últimos a Nagel (1970), McDowell (1978) y Platts (1979).

⁶⁵ Estos elementos, desde luego, son relevantes para establecer el valor de verdad de la razón normativa correspondiente.

así levanto el vaso y bebo del líquido, siempre sería relevante mencionar el aspecto *motivador-explicativo* detrás de mi acción: ‘quería beber agua y creía que el vaso contenía agua’.⁶⁶ Esta *explicación* de mi acción es independiente conceptual y causalmente de la razón normativa correspondiente, pues es falso que ‘sea deseable para mí beber agua de ese vaso’. De otro modo: el sentido motivacional y el sentido normativo de una razón para actuar están separados conceptual y psicológicamente. Ahora bien, según Smith, un problema para la teoría de la motivación de Hume es que sólo parece ocupada con el sentido de las razones motivadoras y desconoce el sentido normativo (cf. Smith 1994, p. 98). La preocupación de Smith por ofrecer una teoría de la motivación susceptible a crítica racional lo conducen a resolver este problema desarrollando una ‘teoría anti-humeana de las razones normativas’ (cf. Smith 1994, cap. 5).

Smith introduce una nueva distinción para explicar a la acción intencional (la cual es el objeto o fin de una razón para actuar). De acuerdo con él, la acción intencional se puede explicar desde dos perspectivas, la intencional y la deliberativa (presenté antes esta última, como la estrategia equivalente a la interpretación de la teoría del observador ideal de Hume). Mediante la perspectiva intencional explicamos a la acción intencional «citando los complejos estados psicológicos que la producen, es decir, ajustamos a la acción intencional dentro de patrones de explicaciones teleológicas y causales». Ahora bien, en atención a la distinción entre los sentidos motivacional y normativo de una razón para actuar, la explicación intencional se explicaría en términos del sentido motivacional (cf. Smith 1994, 131). Por contraste, asumimos una perspectiva deliberativa cuando «explicamos a la acción intencional en términos de patrones de deliberación racional que pudieron o no producir dicha acción». Correspondientemente, esta perspectiva explicativa estaría constituida por el sentido normativo de las razones para actuar. La idea, al parecer, es explicar mediante un proceso deliberativo racional la justificación a favor de la verdad de una cierta razón para actuar. Ahora bien, no es necesario, ni parece que sea alguna vez el caso, que los agentes se involucren conscientemente en tales procesos deliberativos racionales, pero podemos reconstruir los supuestos patrones reflexión racional *ex post facto*. Es decir, podemos explicar ciertas acciones intencionales de los agentes, *una vez que han sido realizadas*, mediante una re-construcción de un proceso racional que adopta la forma general ‘como si’ [*as if*] tal proceso deliberativo racional hubiera tenido lugar y explicará a la acción intencional (cf. Smith

⁶⁶ De acuerdo con Williams, un elemento del conjunto motivacional subjetivo de un agente que dependa de creencias falsas no puede constituir una razón para actuar (cf. Williams 1993, 133).

1994, 131-132). Ahora bien, la distinción entre estas dos perspectivas intencionales expresa, en una nueva dimensión, el problema de cómo se relacionan los deseos y los estados racionales. Smith, en efecto, señala que esta distinción implica una diferencia en la categoría de cada una de las dos explicaciones (cf. Smith 1994, 132). La explicación intencional se corresponde con los deseos, *i.e.*, estados psicológicos genuinos de los agentes relacionados con ciertas acciones, mientras que la explicación deliberativa es una reconstrucción racional de un proceso deliberativo, tal *como si* éste proceso hubiera tenido lugar y fuera relevante citarlo para explicar cierta acción del agente. Lo que Smith espera de esta estrategia es que la perspectiva deliberativa logre influir eficazmente sobre la perspectiva intencional, de manera que los agentes logren generar los motivos para actuar consistentes con la perspectiva deliberativa.

Lo que Smith espera tanto del sentido normativo de una razón para actuar como de la explicación deliberativa es la posibilidad de introducir el hecho de que, *ocasionalmente*, los agentes evalúan lo que pretendemos hacer o lo que de hecho hacemos, es decir, evaluamos la deseabilidad de lo que pretenden hacer. Hemos visto, en efecto, que hay una brecha entre tener una razón motivadora para hacer ϕ y tener una razón normativa para hacer ϕ . Pero si es verdad para un agente que existe una razón para que haga ϕ (en sentido normativo), tenemos que responder, a continuación, “cómo es que aceptar una razón normativa puede ser tanto *sujetarse a* ciertos deseos y no obstante *mantenerse a parte* de tener deseos” (cf. Smith 1994, 136). La primera condición, *sujetarse a ciertos deseos*, proviene del sello humeano de la teoría de la motivación al que Smith pretende circunscribirse. La segunda condición nos dará la respuesta al problema. Ahora bien, según Smith solo podemos entender la peculiar actitud de evaluar que estamos presentando bajo dos únicas soluciones posibles: Evaluar o es un asunto de desear o es un asunto de creer. Smith se propone defender la segunda posibilidad (cf. Smith 1994, 137). Si esta defensa es plausible, Smith sostiene que alcanzará la *conexión conceptual* necesaria para explicar cómo las creencias se relacionan con los deseos, precisamente en el sentido de la racionalidad práctica que interesa defender, de acuerdo con la siguiente condición:

Si un agente cree que tiene una razón normativa para ϕ , entonces ella racionalmente debería desear hacer ϕ (Smith 1994, 148).

Esta condición garantizaría la racionalidad de los procesos deliberativos y explicaría, conceptualmente, cómo se relacionan creencias y deseos. La racionalidad de esta relación la marcaría el modo en que la creencia en una razón normativa afectaría al deseo correspondiente

con la acción. Así, el argumento de Smith tiene que responder cómo es que una razón normativa consiste en una evaluación y cómo es que dicha evaluación consiste en una creencia. La conexión conceptual que busca Smith, con la que pretende solucionar el problema moral, depende de esta respuesta.

En su defensa, Smith recurre a uno de sus recursos de análisis más poderosos en *The Moral Problem*: ofrecer un recuento y una valoración de los lugares comunes [*platitudes*] en torno a una noción, en este caso, la noción de ‘razón normativa’. Al recurrir a estos lugares comunes establece ciertos sentidos básicos que comúnmente reconocemos como constitutivos del sentido de la noción en cuestión. Posteriormente, mediante análisis y reflexión, Smith procura «refinar nuestra comprensión de tales nociones». En la noción de ‘razón normativa’ encuentra que el lugar común está en torno a nociones relacionadas con el *consejo*: “el lugar común está relacionado con un todo de lugares comunes anfitriones del *consejo* [*advice*]. Si tú no estás seguro de qué hacer en alguna situación, ¿cómo deberías proceder para decidir qué hacer? La respuesta es que deberías sacar provecho de la sabiduría de la gente; deberías pedir consejo. Deberías consultar a alguien mejor situado que tú mismo respecto de saber qué deberías hacer, alguien te conozca bien” (Smith 1994, 151). En la situación presentada el agente encara alguna pregunta práctica del tipo ‘¿qué debo hacer en esta situación?’ Lo que despierta al proceso deliberativo es una situación inicial en la que hay una duda práctica (reflejada en la pregunta mencionada) y esto coloca al agente en posición de pedir consejo o guía para responder a la pregunta. Ahora bien, según Smith, este lugar común desemboca naturalmente en la idea de que, adecuadamente idealizado, cada una o cada uno de nosotros es la mejor persona para aconsejarse a sí misma o a sí mismo. De aquí deriva el sentido normativo-racional que necesita su argumento.

Según Smith, «el lugar común del consejo es indicarnos que lo deseable para nosotros es hacer aquello que desearíamos si fuéramos completamente racionales» (cf. Smith 1994, 151). A partir de este punto Smith pasa del lugar común al análisis. El desarrollo del mismo nos enseña que mediante la razón normativa evaluamos nuestra acción desde dos planos. Un plano, en el que evaluamos y planeamos la acción que pensamos realizar, Smith lo llama ‘el mundo posible evaluador’. En este mundo consideramos lo que sería deseable hacer para nosotros *si fuéramos* completamente racionales y no como de hecho somos. ‘Como de hecho somos’ sucede en el segundo plano, el plano al que llama ‘mundo posible evaluado’. En este mundo actuamos realmente, ahí suceden nuestras acciones, y es el mundo que evaluamos. De manera que instalados

en el mundo posible evaluador consideramos las circunstancias de nuestros sucesos en el mundo posible evaluado. Nuestro yo evaluador es, según Smith, plenamente racional y tiene el mejor punto de vista sobre nuestro yo evaluado. Deliberamos sobre qué deberíamos hacer bajo la perspectiva plenamente racional a la que podemos acceder desde el mundo posible evaluador (cf. Smith 1994, 151). La conexión conceptual entre razones normativas y motivadoras parece ahora cerrarse.

Según su análisis, “una creencia evaluativa [o una razón normativa] es simplemente una creencia acerca de qué podríamos desear si fuéramos completamente racionales”. La conexión con los deseos viene, precisamente, de que en la construcción de esta noción de creencia evaluativa hemos «aceptado que se requiere que seamos racionales» (Smith 1994, 160). Es decir, un agente deliberativo smithneano, en la medida en que acepta el requisito de adoptar una perspectiva plenamente racional sobre el contenido de sus deliberaciones, acepta los deseos que tal agente tendría, precisamente en virtud de adoptar tal punto de vista plenamente racional. La condición es incluso aceptable si la formulamos prescindiendo de la noción de creencia, de tal manera que sea claro cómo es que este requisito de plena racionalidad conduce a ciertos deseos prácticos: “Nuestro hacer ϕ es deseable sólo en el caso en que desearíamos hacer ϕ si fuéramos completamente racionales” (Smith 1994, 177). Si este argumento es correcto, entonces Smith ha logrado establecer que existe en efecto una conexión conceptual entre el deseo y la razón. A continuación ofreceré una revisión de la teoría de Hume, aceptando la plausibilidad de una explicación instrumental del papel de la razón.

4. La teoría de los juicios morales de Hume

En este capítulo analizo la teoría de los juicios morales de Hume. En primer lugar, presento brevemente los argumentos por los cuales sostiene que los juicios morales se basan en sentimientos, de modo que no son objeto de la razón. Después, examino dos interpretaciones dominantes de esta idea: la sentimentalista y la del observador ideal. La primera acepta literalmente que los juicios morales son la expresión de un sentimiento, pero enfrenta problemas cuando trata de explicar a qué se refiere dicho sentimiento. Examino estas dificultades. Defiendo que la interpretación del observador ideal no enfrenta estos problemas porque, con base en la idea de que lo distintivo del sentimiento moral es que está basado en un punto de vista general o imparcial sobre el carácter de una persona (que puede ser el propio agente), no tiene que ocuparse de explicar a qué se refiere tal sentimiento. En efecto, se refieren al carácter de una persona, pero éste es considerado idealmente, esto es, desde un punto de vista general e imparcial, de modo que su referencia es, por decirlo así, un modo de sentir o experimentar a alguien. Ésta es la posición que defiendo en este trabajo.

4.1. Argumentos con los que Hume prueba que las distinciones morales no se derivan de la razón

La siguiente pregunta de Hume expresa sus dos intereses principales respecto de la moralidad: “¿de qué naturaleza son estas impresiones [las morales] y de qué modo actúan sobre nosotros?” (T 3.1.2.2 SB 470). En efecto, con su investigación sobre la moralidad Hume quiere establecer en primer lugar la naturaleza de nuestras percepciones morales y, en segundo lugar, como tales percepciones afectan a nuestra voluntad o intervienen en la motivación. Ahora bien, dado su modelo de la mente, para él el pensamiento moral tiene que ser o bien una impresión o bien una idea. Por ello, el tercero y último libro del *Tratado*, titulado “De la moral”, inicia con la discusión en torno a esta pregunta: *¿Distinguimos entre vicio y virtud, y juzgamos que una acción es censurable o digna de elogio, por medio de nuestras ideas o de nuestras impresiones...?* (T. 3.1.1.2 SB 456). La pregunta presupone que de hecho hacemos distinciones morales e inicia la investigación al respecto sin preguntar, previamente, por el *origen* de términos morales como virtud, vicio, motivo, deber, meritorio, benevolencia, justicia, etc. (cf. Fate Norton 2003, 13). Nuestras distinciones morales presuponen tales términos –y las ideas correspondientes-, pero

Hume parece sugerir que nuestro uso de ellos no representa una dificultad de interés para distinguir entre el vicio y la virtud. Lo que le interesa es estudiar la diferencia entre el vicio y la virtud señalando cuál es el origen y las características distintivas de nuestros *juicios* sobre lo moralmente correcto y lo moralmente incorrecto (cf. Stroud, 251-252). Así, eventualmente, una dificultad para su investigación será explicar cómo las impresiones morales dan origen a un tipo de juicio porque, como veremos en seguida, para Hume los pensamientos morales con los que distinguimos entre vicio y virtud descansan en sentimientos.

Hume argumenta que si la distinción entre vicio y virtud se originara en nuestras ideas, entonces la razón –la facultad cognitiva– podría mediante sus operaciones ser el fundamento de ella (T 3.1.1.4 SB 456-457). En consecuencia, los juicios morales serían el producto de la razón y un examen de sus propiedades no nos dejaría dudas sobre la pertinencia de la razón para fundamentar a la moralidad. Tales propiedades son enlistadas en el siguiente pasaje, el cual parece caracterizar a una posición racionalista sobre la moralidad, enumerando las diferentes tesis que comprendería:

...que existe en las cosas una eterna adecuación o inadecuación y que ésta es idéntica para todo ser racional que la contemple, que las medidas inmutables de lo justo y lo injusto imponen una obligación no solamente a las criaturas humanas, sino hasta la misma Divinidad; quienes dicen todas estas cosas sostienen unos sistemas que coinciden en afirmar que la moralidad, como la verdad, se discierne por medio de ideas, por su yuxtaposición y comparación (T 3.1.1.4 SB 456-457).

Estas propiedades parecen estar relacionadas entre sí e impondrían sobre los agentes restricciones psicológicas para pensar y actuar en conformidad con los juicios morales (aunque no es claro cuál es el origen de tal imposición u obligación, especialmente para la Divinidad, porque reconocer lo justo o lo injusto, o lo verdadero o lo falso, no produce, necesariamente, una obligación para actuar en consecuencia).⁶⁷ Por otra parte, la afirmación de que “existe en las cosas una eterna adecuación o inadecuación” presupone que las relaciones morales suceden de acuerdo con una cierta regla o regularidad moral que es instanciada o reflejada “por las cosas” e inteligible mediante la razón por medio de una idea o juicio que representaría tales relaciones. La propiedad

⁶⁷ Con su caracterización de una posición racionalista sobre la moralidad Hume parece decir que las propias relaciones morales establecidas racionalmente *generan* las obligaciones o mandatos respectivos para actuar de acuerdo con ellas, so pena de irracionalidad. Ésta es una idea que parece anticipar las ideas de Kant al respecto, sin embargo, como Smith ha señalado, explicar cómo una consideración racional logra producir un motivo para actuar es precisamente el problema moral por excelencia (cf. Smith 1994, §1.3). Desarrollaré más adelante esta cuestión. A favor de Hume, por ahora, sólo recordaré que él precisamente está rechazando esta posición y argumentará que las consideraciones de la razón por sí solas no producen ninguna obligación ni disposición para actuar al respecto.

de la objetividad se derivaría de que tal adecuación es “idéntica para todo ser racional”, incluyendo a Dios. Además, el carácter normativo de tales reglas valdría universalmente para todo ser racional e impondría obligaciones específicas para actuar en conformidad con ellas. Por último, “su yuxtaposición y comparación” parecen propiedades lógicas de composición, de manera que podemos comparar, combinar y relacionar lógicamente los juicios morales sin pérdida de su verdad, la cual, a su vez, consistiría en la “eterna adecuación” de la regla con las cosas.⁶⁸ Quien sostuviera esta posición, caracterizada de este modo, sostendría al mismo tiempo una concepción racionalista sobre la moralidad. Hume rechaza esta posición y, como veremos en seguida, argumenta en contra de que las distinciones morales se fundamenten en la razón. Dados los términos en que Hume ha establecido el problema inicial de la moralidad, si su argumentación es correcta, entonces se seguiría de ella que «las distinciones morales se derivan de un sentimiento moral» y no de la razón (3.1.1.9-10 SB 458). A continuación examinaré los argumentos por los cuales rechaza que el fundamento de los juicios morales sea la razón y, posteriormente, examinaré qué propiedades caracterizan a dicho sentimiento moral.

En T 3.1.1. Hume desarrolla seis argumentos para probar que la razón no fundamenta a la moralidad. En una paráfrasis resumida, éstos son:

- 1º **3.1.1.4-8 SB 456-458:** 1. La moral influye en las acciones y afecciones (porque pertenece a la esfera *práctica* de la filosofía, no a la *especulativa*. Hume divide, en este pasaje, a la filosofía solo en estas dos esferas). 2. La sola razón no puede influir sobre las acciones porque es especulativa: procede buscando adecuaciones en las cosas, p.e., suponiendo que lo justo o lo injusto son medidas inmutables e idénticas para todo ser racional. Si fuera así, de otro modo, la razón podría descubrir el bien o el mal mediante una deducción racional (cf. 2.3.3). Por lo que se sigue que *la moralidad no podrá derivarse de la razón*.⁶⁹
- 2º **3.1.1.9-10 SB 458:** 1. La razón consiste en el descubrimiento de la verdad o falsedad, y éstas consisten en un acuerdo o desacuerdo con relaciones reales de ideas, o de éstas con la existencia y los hechos reales. La razón es inactiva. 2. Las pasiones, voliciones y acciones son incapaces de tal acuerdo o desacuerdo, porque son hechos y realidades originales, completos en sí mismos, sin implicar referencia alguna a otras pasiones, voliciones o

⁶⁸ Aunque Hume no menciona a ningún «racionalista» por su nombre, parece claro que con este pasaje se refiere a ideas sostenidas por los racionalistas de su tiempo, como R. Cudworth (1617-1688), W. Wollaston (1659-1724) y S. Clarke (1675-1729). Los editores de la edición de Oxford del *Tratado*, D. Fate Norton y M. J. Norton, comentando este pasaje, indican que Locke argumenta que con base en la conformidad o disconformidad de las acciones con las reglas derivamos nuestras ideas sobre el bien y el mal morales, aunque él no dice que tales reglas sean eternas e inmutables (remiten a Locke *Essay* 2.28. 14-15).

⁶⁹ En otras palabras, el argumento prueba que la virtud no puede ser objeto de la razón, porque no puede ser conforme o mesurable como lo son las ideas de la razón (cf. Karlsson, 237). Quiero enfatizar la importancia de este argumento para el debate entre cognitivistas y no-cognitivistas. Con base en la distinción de Hume entre la esfera «práctica» y la esfera «especulativa» de la filosofía, Hume ha confina a la razón a tener un papel inerte o pasivo para la vida práctica. Aceptar o rechazar a esta idea marca al mismo tiempo la diferencia entre los cognitivistas y los no-cognitivistas, *a pesar* de que Hume no parece argumentar sobre la distinción ni admitir, siquiera, alguna alternativa. De este modo, este argumento negativo en contra de la razón como fundamento de la moralidad apoya sólo por exclusión, *sin más*, la tesis positiva de que moralidad *es* un asunto de los sentimientos (cf. Botros 2006, 33).

acciones 3. Las acciones no son razonables o irrazonables, son laudables o censurables, meritorias o demeritorias. Luego, *las distinciones morales no son producto de la razón*.⁷⁰

- 3° **3.1.1.11, 12, 14 SB 459-460 (cf. 2.3.3. 5-10 SB 415-418)**: 1. La razón podría influir oblicuamente sobre las pasiones (i) informándonos de la existencia de un objeto deseado o (ii) descubriendo la conexión causal para obtener un fin. 2. Solemos tener creencias tipo (i) y (ii) equivocadas, pero no decimos que tales errores son las fuentes de la inmoralidad. Sólo son errores de hecho. 3. Las acciones no producen jamás juicio alguno. En suma, *la distinción entre el bien y el mal no puede ser hecha por la razón porque ella influye sobre la acción y la razón es incapaz de ello*.⁷¹
- 4° **3.1.1.13 SB 460**: 1. Las distinciones de la razón no admiten variación en grado, ni la verdad ni la falsedad ni ninguna otra distinción. 2. Las distinciones morales, como el vicio o la virtud, si admiten grados. (No nos parece igual de grave un robo por hambre que un robo de cuello blanco, ni un asesinato accidental que uno alevoso.) Por lo tanto, *el fundamento de las distinciones morales no es la razón*.
- 5° **3.1.1.17-26 SB 463-469**: Si la virtud fuera descubierta por el entendimiento tendría que ser objeto de la comparación de ideas o de una inferencia sobre hechos (no hay otras operaciones de la razón o entendimiento). Las únicas relaciones entre ideas operadas por la razón son: Semejanza, contrariedad, grados de cualidad y proporciones en cantidad y número. Por tanto, la moralidad (el vicio o la virtud) no se encuentra en ninguna de estas relaciones. Ilustraciones: El vástago de un olmo al ir creciendo acaba y destruye al árbol padre. Las relaciones son las mismas que en el parricidio –aunque por causas distintivas– pero es claro que ellas no son suficientes para calificar a esta relación como inmoral (cf. §24) Observamos algo similar respecto del incesto "cuando exactamente la misma acción y las mismas relaciones en los animales no presentan la menor depravación

⁷⁰ Con base en este pasaje, Baille atribuye a Hume una posición internista sobre la motivación. De acuerdo con él, Hume sostendría que los juicios morales, aunque no son producto de la razón, tienen la capacidad de influir sobre la razón (cf. Baille 123). Pero esto no parece correcto porque ante un mismo juicio moral un agente podría estar motivado para actuar conforme con él y otro podría no estarlo. Éste último agente carecería del deseo apropiado para actuar de acuerdo con el juicio moral. De este modo, la relación de los juicios morales, en cuanto tales, con un deseo apropiado para despertar la acción es contingente, es decir, no es interna al juicio. De modo que la posición de Hume puede describirse mejor como externista. La posición racionalista que Hume ha presentado antes, según la cual los juicios morales imponen restricciones psicológicas obligatorias de manera universal –para cualquier ser racional– sí sería un ejemplo de internismo, porque según ella es propio o inherente a los juicios morales motivar acciones determinadas. Pero Hume, precisamente, argumenta contra esta posición racionalista.

Más interesante, quizá, es decidir cómo el juicio moral puede suscitar o relacionarse con un deseo para constituir un motivo para actuar. Es posible que los juicios morales puedan inducir, por sí mismos, un cierto deseo, aunque éste no sea la causa inmediata de la acción. Por contraste, parece que los juicios o creencias comunes de la razón, los que no son morales, parecen carecer de esta capacidad de suscitar deseos (cf. Pigden, 82).

⁷¹ Este argumento trae, quizá, los retos más difíciles para los intérpretes de la teoría de la motivación de Hume. Hampton, por ejemplo, con base en él sostiene que la razón no puede por sí misma movernos a actuar, a menos que concurra con un cierto deseo. De este modo, la razón tiene sobre la acción un mero efecto *causal*, pero no motivacional (cf. Hampton, 63). En otras palabras, entre una razón para actuar y un motivo para actuar no existiría ningún vínculo conceptual, sino solo alguno causal. Una posición semejante es sostenida por Karlsson. De acuerdo con este pasaje, él sostiene que “estas cogniciones nos *causan* formar un deseo de actuar de cierta manera para obtener el placer que un objeto tiene en reserva”. Pero no existe ningún otro vínculo entre la razón y la motivación. La razón, dice él, se ocupa de la verdad y de la falsedad, no de los motivos para actuar. Así, el motivo para actuar reside en las pasiones, las cuales son inherentemente motivacionales, no contingentemente (cf. Karlsson 248-249). Por lo que he dicho antes sobre la relación entre creencias y pasiones creo que esta posición no refleja correctamente la visión general de Hume sobre la motivación (v. supra §3.1).

Por otra parte, de la afirmación de Hume de que los errores prácticos por creer que existe un cierto objeto o por considerar suficientes a ciertos medios para alcanzar un fin son errores de hecho, pero no errores morales, se deriva que los agentes morales usualmente actúan motivados bajo la apariencia del bien [*under the guise of the good*], esto es, que aun cuando se trate de un error de hecho, el agente moral mantiene una perspectiva de los asuntos en cuestión como si se tratara de lo moralmente correcto, de lo que tiene que hacerse según el bien moral, etc. Esta consideración amplía la brecha que hay entre la razón y los juicios morales, porque las consideraciones racionales sobre hechos no son tan eficaces, para despertar un motivo, como lo es por sí mismo la apariencia del bien, por equivocada que ésta pueda ser.

ni fealdad morales" (cf. §24). La experiencia enseña que ni algunas cuestiones de hecho son descubiertas por la razón ni derivadas de alguna relación distintiva entre ideas. Por ejemplo, el caso de una acción reconocida como viciosa. Consideremos el asesinato intencionado. Vamos a examinarlo desde todos los puntos de vista posibles, a ver si podemos encontrar esa cuestión de hecho o existencia que llamamos vicio. Nunca podremos descubrirlo, sino hasta el momento en que dirijamos la reflexión a nuestro propio pecho y encontremos allí un sentimiento de desaprobación que en vosotros se levanta contra esa acción. Ésta es una cuestión de hecho, pero es objeto del sentimiento, no de la razón. Por consiguiente: *El vicio y la virtud pueden compararse con los sonidos, colores, calor y frío, que, según la moderna filosofía, no son cualidades en los objetos, sino percepciones en la mente.*⁷²

6° 3.1.1.27 SB 469-470: Toda teoría moral da una explicación deficiente de la obligación moral. En algún momento el autor en cuestión cambia la cópula habitual "es" por "debe" o "no debe", con lo que establece una nueva relación o afirmación, pero no da ninguna razón de cómo es posible que esa nueva relación se deduzca de otras totalmente diferentes. En otras palabras, los autores morales no ofrecen ninguna deducción aceptable para inferir de enunciados descriptivos enunciados valorativos o normativos. En consecuencia, *la razón no explica a los enunciados valorativos o normativos.*⁷³

Si bien estos argumentos no prueban positivamente que la distinción moral entre vicio y virtud consista en un tipo de sentimiento, dados los términos en que Hume ha propuesto el problema (cf. T 3.1.1.2 SB 456), se sigue ahora, por vía negativa, que «la exclusión de una, las ideas, sea un argumento convincente para la otra, las impresiones» (cf. T 3.1.2.1 SB 470), esto es, como los argumentos anteriores no han probado que las distinciones morales sean objeto de la razón en

⁷² Por su generalidad y abstracción, de estos seis argumentos defendidos por Hume éste es quizá el más difícil de entender (cf. Stroud 154). Por una parte, Hume no argumenta *en ningún lado* por qué todas las relaciones entre ideas tienen que consistir en los cuatro tipos que menciona (v. supra §1.1). Su estrategia es retar a que se ofrezca una alternativa plausible o mejor. De acuerdo con él, si el vicio o la virtud son objeto de la razón, tendría que poder mostrarse que se ajustan al menos a alguna de dichas relaciones entre ideas. Su propia estrategia es ofrecer la analogía con casos que claramente no son morales, a pesar de que las relaciones entre ideas y objetos es semejante y comparable a casos (el parricidio y el incesto) que nos resultan aborreciblemente morales. Como las mismas relaciones no son suficientes para señalar relaciones inmorales en los ejemplos con los que establece la analogía, se descarta que el vicio o la inmoralidad se deriven de las relaciones consideradas. Pero la inmoralidad es manifiesta en esos casos, de modo que su origen no puede ser la razón.

Por otra parte, como el juicio de que el incesto o el parricidio son inmorales no descansa meramente en las relaciones entre objetos, pero sí, *de alguna manera*, en los objetos y en nuestra percepción de ellos, Hume concluye que esta relación puede explicarse también con analogía con las propiedades secundarias de los objetos físicos, por ejemplo, los colores. La percepción de un color no se refiere a una propiedad primaria o intrínseca de un objeto, sino a una cierta condición de nuestra propia percepción que *reacciona* de alguna manera a ciertas propiedades de los objetos. Así, el objeto no es de un cierto color, pero es causalmente necesario para la percepción de un color específico. El vicio o la virtud serían de este tipo de propiedad: no son parte de los objetos externos a la mente, pero tales objetos son necesarios causalmente para nuestra percepción moral (más adelante, al examinar cómo los juicios morales son para Hume un sentimiento, discutiré con profundidad la cuestión de cómo, si acaso, se refieren a propiedades secundarias, v. infra p. 128 y ss). Por ahora, quiero introducir otra dificultad de interpretación.

Fate Norton ha observado que esta posición de Hume, correctamente interpretada, hace del vicio y de la virtud no exactamente *impresiones* secundarias, sino terciarias. En efecto, la percepción de los objetos necesarios para el juicio moral equivaldría a la impresión primaria, de la cual se genera en nuestra mente una *impresión secundaria* o impresión de reflexión, esto es, una idea o ideas de los objetos. Por ejemplo, veo que A se dirige a B y lo apuñala mortalmente. Pero como sé que A es hijo de B, caracterizo a esa relación como parricidio y experimento dolor o desagrado. Si ahora, además, caracterizo a la relación como un vicio y la juzgo como moralmente reprochable, ésta nueva idea es una *impresión terciaria* respecto de la anterior, de modo que es plausible que las ideas de vicio y virtud "puedan ser precisamente tales ideas terciarias" (cf. Fate Norton 2000, 119-20).

⁷³ Este pasaje es célebre por introducir la cuestión es/debe (*the is-ought question*), erróneamente llamada como 'la falacia naturalista'. La supuesta falacia consistiría en que no puede derivarse *lógicamente* un enunciado normativo (con verbo principal debe) de un enunciado descriptivo (con verbo principal es).

tanto que no son ideas, entonces tienen que derivarse, por exclusión, en un tipo de impresión, de ahí que sean un sentimiento o sensación. Antes de considerar esta última idea me quiero detener un poco más en los seis argumentos anteriores y las relaciones que guardan entre sí.

Mackie ha señalado que estos argumentos están acribillados por ambigüedades porque no es claro qué significado de ‘razón’ usa Hume cuando dice que la moralidad no está basada en ella y porque tampoco está claro cómo vamos a interpretar ‘sentimiento moral’ ni cómo las distinciones morales se deben a él (cf. Mackie 1980, 51), pero creo que esta acusación es injusta. Me parece que en estos argumentos se expone detalladamente la concepción que Hume tiene de la razón y cómo se relaciona con los motivos morales. Del sentimiento moral me ocuparé en la siguiente sección. En el argumento (1) indica que la razón no es práctica, en el (2) que es inerte o “inactiva” para producir motivos, en el (4) que sus objetos, esto es, las ideas, creencias o enunciados, en tanto que pueden ser calificados como verdaderos o falsos, no admiten grados (en este sentido un enunciado o creencia no es más o menos verdadero, sino llanamente verdadero o falso, y tampoco ambos) y en el (6) reconoce que no conoce ningún buen argumento racional para justificar enunciados normativos u obligaciones morales. Este grupo de argumentos me parece entre sí consistente. Todos ellos expresan la concepción de que el objeto de la razón pertenece al mundo de las ideas o de la teoría y deriva las implicaciones de tal posición. Al mismo tiempo, mantienen la brecha conceptual que según Hume hay entre las facultades cognitivas y las no-cognitivas. Con todo, podría parecer que este grupo, formado por los argumentos (1), (2), (4) y (6), es inconsistente con el argumento (3), según el cual la razón tiene un papel instrumental, esto es, que tal vez no pertenezca exclusivamente al mundo de las ideas sino que podría tener injerencia en el mundo de la voluntad. Así, una vez más, la relación de (3) con (2) constituye la dificultad central de la teoría de la motivación de Hume de la que pretendo ocuparme en este capítulo.

De acuerdo con (3), la razón tendría alguna influencia sobre la voluntad, así sea en el sentido mínimo de indicar si se han elegido los medios adecuados para alcanzar un fin o si el fin mismo es alcanzable, pero incluso así esquematizado (3) es compatible con (2), toda vez que este argumento señala que la razón es inactiva para motivar acciones, de modo que su papel instrumental solo se activaría bajo el impulso de un deseo o pasión. De modo que según la relación de (3) con (2) –y no parece que haya problema para extrapolar esta idea a los argumentos (1) y (4)- la teoría de la motivación de Hume sostendría que la razón tiene un papel subordinado a las facultades no-cognitivas, un papel meramente instrumental, relativo a los medios de la acción o

de los pensamientos y nunca a los fines. El argumento (6) no parece representar ningún problema de consistencia.

A pesar de que el argumento (6) contiene uno de los pasajes más célebres de Hume (el relativo al es-debe), en realidad Hume mismo no vuelve a este tema en sus obras posteriores – p.e., en la IPM no hace ninguna mención al respecto. Yo sólo voy a tomar la interpretación de Stroud, la cual me parece que permite apreciar fácilmente la consistencia con los otros argumentos que examino. De acuerdo con él, en este pasaje a Hume lo que “parece interesarle ante todo es enfatizar una vez más la tesis de que el hecho de que los juicios morales no puedan ser “percibidos mediante la razón” se debe a su carácter especial”, esto es, que como no son exactamente ideas no pueden ser objeto de la razón (cf. Stroud, 273).⁷⁴ En esta línea, el argumento 6 es perfectamente compatible con (1), (2) y (4) de manera que no agregaría ningún contenido a lo afirmado en éstos. En consecuencia, por transitividad, mantendría la misma relación que éstos tienen con (3), esto es, que en su papel instrumental la razón está subordinada a las facultades no-cognitivas y que en su papel teórico es motivacionalmente inerte. De manera que, en conjunto, parece injusto señalar que la argumentación de Hume es inconsistente, como acusó Mackie. Consideremos ahora la estrategia por la que Hume sostiene que las distinciones morales *se derivan de un sentimiento moral* y las características de tal sentimiento.

4.2. La interpretación *sentimentalista* del sentimiento moral

Según el modelo de la mente de Hume y en tanto que objetos de ella, las distinciones morales que hemos considerado hasta ahora se basan en impresiones o ideas. La argumentación examinada en la sección anterior excluye que sean ideas –o podrían ser objeto de la razón- (“la exclusión de las unas constituye un convincente argumento a favor de las otras”⁷⁵). De estas consideraciones se desprende la tesis principal de Hume sobre la moralidad: “La moralidad es, pues, más propiamente sentida que juzgada, a pesar de que esta sensación sea por lo común tan débil y suave que nos

⁷⁴ Stroud, en el mismo lugar, dice también que este pasaje fue una ocurrencia de último momento que Hume incluyó, creyendo que le sería útil más adelante.

⁷⁵ Desarrollé antes (v. supra p. 122 y ss) que la argumentación positiva de Hume a favor de entender a la moralidad como un sentimiento descansa casi exclusivamente en la exclusión de la razón como fundamento de la moralidad. Creo que una presentación no excesivamente caritativa de la estrategia de Hume tendría que reconocer que él argumenta adoptando una perspectiva sistemática del *Tratado*, esto es, que los descubrimientos que presenta en el Libro I sobre la estructura de la mente y sus materiales básicos (las impresiones y las ideas) ahora son relevantes para abordar el tema de la moralidad. En conjunto, su estrategia favorece la interpretación general del *Tratado* como si se tratará de un solo argumento.

inclinemos a confundirla con una idea...” (cf. T 3.1.2.1 SB 470). La interpretación sentimentalista acepta literalmente esta afirmación. Con todo, tenemos que considerar aún una segunda característica relativa a la fuerza motivacional de las distinciones morales: “...la impresión surgida de la virtud es algo agradable, y... la procedente del vicio es desagradable” (cf. T 3.1.2.2 SB 470). De este modo, las propiedades asociadas con el placer y el dolor se comunican al vicio y a la virtud con efectos similares: procuramos al placer o virtud, y nos alejamos del dolor o vicio (v. capítulo anterior y T 3.3.1.2 SB 574). Según esto, existe una fuerte determinación natural que nos motiva según la naturaleza fenoménica de las percepciones morales. Hume respalda a esta idea «en la experiencia»:

No existe espectáculo tan hermoso como el de una acción noble y generosa, ni otro que nos cause mayor repugnancia que el de una acción cruel y desleal. No hay placer comparable a la satisfacción que nos proporciona la compañía de aquellos a quienes amamos y apreciamos, de igual modo que no hay tampoco mayor castigo que el verse obligado a vivir con quienes odiamos o despreciamos (T 3.1.2.2 SB 470).

Las acciones que nos parecen nobles, generosas, crueles o desleales, nos provocan reacciones similares de agrado o placer comparables a la compañía de las personas que amamos o de desagrado y repugnancia de quienes odiamos. El defensor de la interpretación sentimentalista tiene lidiar ahora con el pensamiento de que respuestas emotivas parecen *referirse* a las *cualidades* de ciertas personas, por las cuales nos resultan amables u odiosas. Esto parece inconsistente con la idea de que la moralidad surge de sentimientos del observador, porque aquí Hume parece sostener que éstas son propiedades morales propias o inherentes de ciertos objetos, ciertas personas, de las cuales tenemos una percepción emotiva por medio del sentimiento moral. Por ejemplo, una persona generosamente brinda su ayuda a otra en desgracia y por medio del sentimiento moral percibimos a la generosidad como un rasgo moral propio de esa persona. El sentimiento moral se experimenta como un cierto placer provocado por la apreciación de la generosidad de la persona. Por estas consideraciones, Stroud afirma que estas ideas implican “que la virtud y el vicio son rasgos *objetivos* de una acción o una cualidad personal; son los rasgos que hacen que tengamos ciertos sentimientos cuando observamos o contemplamos dicha acción o dicha cualidad” (cf. Stroud, 266). Stroud nos remite a un pasaje de EPM donde, de acuerdo con él, Hume sostendría que “...que la virtud de una acción es una propiedad objetiva de ésta”, porque «define» ahí a la virtud como “*toda acción o cualidad que da al espectador el placentero sentimiento de la aprobación*” (EPM Ap. 1, SB 289; Stroud, 266). La interpretación sentimentalista, al menos en esta lectura de Stroud, tiene que aceptar esta implicación. Una

solución a esta inconsistencia proviene de la propuesta interpretativa de Fate Norton a través de lo que llama la «teoría *causal* de la percepción moral» (Fate Norton 2003, 163-164).

El núcleo de dicha teoría es la idea de imparcialidad. Según Hume, la marca distintiva del sentimiento moral es la imparcialidad: «cuando el carácter de otra persona, incluso un enemigo, es considerado en general y sin referencia a nuestro interés particular causa esa sensación o sentimiento en virtud del cual lo denominamos moralmente bueno o malo» (cf. T 3.1.2.4 SB 472; cito extensamente este pasaje más abajo, al examinar la teoría del observador ideal). Fate Norton acepta que la idea de imparcialidad se refiere a la intención, motivo o carácter de los agentes detrás de sus acciones y que son éstos los que hacen surgir los peculiares placer y dolores distintivos del sentimiento moral (cf. Fate Norton 2006, 163-4). Según esta teoría, “las acciones mismas, moralmente hablando, sólo son signos [*signs*] de los motivos que las producen” (cf. Fate Norton 2006, 164). De este modo, un agente *percibe* los motivos por los que otro agente actúa por medio de los signos *externos* de ellos, la acción del agente. Los sentimientos morales peculiares mencionados equivalen al reconocimiento de tales motivos y éste es posible porque descansa en la idea de Hume de que la moralidad tiene «una base estable e inalterable» en la propia naturaleza humana. En combinación con las circunstancias sociales, dicha base aporta rasgos substantivos de la vida moral que se manifiestan en la experiencia moral y en las distinciones morales (cf. Fate Norton 2006, 159-160). La propia sensibilidad moral, la cual incluye por sí misma motivaciones morales relevantes, es parte de este equipamiento natural y su función parece ser la de producir en nosotros sentimientos distintos y peculiares que reflejan dichos motivos morales (cf. Fate Norton 2006, 162). De este modo, al explicar a la valoración moral, Fate Norton reduce el carácter y las acciones de una persona a la cualidad de ser signos de los motivos morales característicos distintivos de la naturaleza humana. De manera que cuando juzgamos a una persona como viciosa, por ejemplo, lo que hacemos es aplicar un cierto término moral –vicio- y la *idea* de vicio a una persona que posee las características correspondientes. Al parecer, es la observación de esas características lo que provoca un sentimiento de desagrado hacia la persona y, en este caso, éste es un sentimiento moral (cf. Fate Norton 2003, 114-15).⁷⁶

⁷⁶ Mi paráfrasis de la explicación del sentimiento moral hecha por Fate Norton no es del todo justa, porque él dice que en el sentimiento moral se aplica “correctamente” el término moral, viciosa en este caso, a un cierto individuo. A mí me parece que aun cuando la idea de Hume sí es establecer cómo hacemos uso de términos valorativos específicamente morales por medio del sentimiento moral, no es claro que siempre los apliquemos correctamente. Así, un recurso teórico que emplea para hablar de cómo tenemos que esforzarnos para aplicar correctamente tales

Según esta interpretación de Fate Norton, las propiedades morales *son* como las propiedades secundarias de los objetos, esto es, no pertenecen intrínsecamente a ellos, pero requieren de ciertas propiedades de los objetos para que las podamos percibir. Hume parece entonces sostener que las propiedades morales son tanto propiedades *secundarias* como propiedades *objetivas* de una acción o de una cualidad personal. Diversos intérpretes han tratado de evitar esta conclusión favoreciendo una interpretación *sentimentalista* no objetivista.⁷⁷

La base sentimentalista de la interpretación se mantiene porque acepta que los juicios morales son la expresión de sentimientos humanos reales (según la caracterización de E. Radcliffe 1994, 37 –aunque hay variaciones en su formulación, por ejemplo, entre otras, Mackie 1980, 64 y 73; Merrill, 40; Botros 2006, 34; Stroud 183 y ss.; y Foot 1978, 79). Con variaciones, la estrategia que han seguido consiste en minimizar las implicaciones objetivistas, pero minimizar no es eliminar. Con todo, a juzgar por la variedad de interpretaciones no parece fácil presentar a la posición sentimentalista y, por lo tanto, tampoco la respuesta a la inconsistencia en cuestión. Presentaré en seguida una breve revisión de tales variaciones sentimentalistas.

Para empezar, Mackie señala que no hay duda sobre cuál es el punto principal de Hume: el hecho esencial es que cuando distinguimos al vicio de la virtud Hume se refiere sencillamente a que la gente tiene diferentes sentimientos respecto de ellos (Mackie 1980, 64). Según Mackie «literalmente» no hay tales características objetivas, sino sólo ciertas sensaciones o sentimientos respecto de las acciones a las que adscribimos virtud o vicio, de manera que las propiedades morales son como propiedades secundarias (cf. p. 73). Pero esta posición todavía hace plausible una interpretación objetiva de las ideas de Hume, porque al menos un aspecto de la virtud o el vicio tendría que ser aquello a lo que los agentes reaccionan sensiblemente: las *acciones* correlativas con los sentimientos morales.

Por su parte, Merrill explica que ‘sentido moral’ [*moral sense*], como se desprende de EPM se refiere a un “sentido o sensación interna” y afirma que esta idea no representa ningún

términos e ideas valorativas es recurrir a un cierto punto de vista autorizado, como veremos, el punto de vista impersonal de alguien «de buen juicio y sentido» (v. abajo, siguiente párrafo). Por esta consideración preferí una paráfrasis más neutral respecto de la corrección del término moral contenido en el sentimiento moral.

⁷⁷ De acuerdo con Botros, D. Fate Norton es el único intérprete notable que defiende que Hume es un objetivista moral (v. Botros 2009, 31; *David Hume, common-sense moralist, sceptical metaphysician* 1982, 116, note 21).

cambio sustancial respecto de la doctrina del *Tratado*. Sin embargo, como vimos, esta observación se responde con la «definición» objetiva de vicio o virtud de EPM Ap. 1, SB 289 citada por Stroud.

El propio Stroud marca una salida de la interpretación objetivista proponiendo que ‘*X* es vicioso’ podría significar que ‘*X* tiene una cualidad –sin importar cuál sea ésta– que en realidad causa que tengamos un cierto sentimiento de desaprobación hacia *X*. Esto impediría que la viciosidad pudiera identificarse esencialmente con alguna propiedad o característica objetiva determinada, y tendría la virtud de enfatizar la importancia de los sentimientos para la existencia misma de los vicios y las virtudes...’ (Stroud, 267). Nuevamente, aunque esto elimina la interpretación objetivista fuerte del vicio o la virtud como propiedades objetivas, nos deja de todos modos con la idea de que hay una propiedad indeterminada en los objetos con la que se corresponde el sentimiento moral.

Por su parte Botros comentando y siguiendo la discusión de Stroud sobre estos pasajes que respaldan una interpretación objetivista, en particular EPM Ap. 1, SB 289, dice que quizás Hume pudo haberse expresado aquí menos «engañosamente» (Botros 2006, 31). De acuerdo con ella, “...la idea general de su análisis es distinguir estas propiedades objetivas de la virtud”. Su interpretación descansa, entonces, en el propósito general de su análisis. No estoy seguro que su posición final disuelva por completo las dificultades en torno a las propiedades objetivas: “Porque aunque estas propiedades son, como el señala en 3.3.1.30, las “fuentes” de las que surge nuestro placer, no hay nada especial por e.g., su utilidad para la sociedad, la cual podría dictar lógicamente que sintamos este placer cuando contemplamos acciones socialmente útiles” (Botros 2006, 31). Yo creo que puede ofrecerse una interpretación plausible de Hume, que no caiga en estas dificultades, desarrollando la interpretación del *observador ideal* de los juicios morales.

4.3. La interpretación del *observador ideal* del sentimiento moral

Tal como la caracteriza Elizabeth Radcliffe, esta posición puede presentarse así: “La lectura del observador ideal, por otra parte, sostiene que, para Hume, las distinciones morales están basadas en los sentimientos hipotéticos de un espectador ideal –quien es, entre otras cosas, alguien completamente informado, enteramente objetivo y no auto-interesado” (Radcliffe 1994, 37). Bajo esta idea, defenderé que las supuestas propiedades objetivas mencionadas arriba son parte de la experiencia normativa y valorativa que un agente tiene de la realidad social al pertenecer a una cierta comunidad e interiorizar sus diferentes estándares normativos.

El sentimiento mismo de agrado o desagrado constituye, para Hume, la alabanza, admiración, rechazo o censura (cf. T 3.1.2.3 SB 471). Hume parece referirse a la experiencia común de la moralidad, en la cual no nos interesamos por la naturaleza ni el origen de las distinciones morales: “No vamos más allá ni nos preguntamos por la causa de la satisfacción. No inferimos la virtud de un carácter porque éste resulte agradable; por el contrario, es al sentir que agrada de un modo peculiar cuando sentimos de hecho que es virtuoso” (T 3.1.2.3 SB 471). Hume señala que es el propio sentimiento de agrado o desagrado el que llama la atención sobre un cierto asunto moral. Experimentamos placer o dolor como una expresión de aprobación o de censura en presencia de una cierta persona y su carácter. Sin este sentimiento, nuestra percepción de la persona, de sus acciones o de sus cualidades carecería de relieve o de textura –como cuando contemplamos el mobiliario de una habitación sin experimentar ninguna valoración o emoción. Esta observación de Hume es consistente con sus argumentos para rechazar que la razón sea fundamento de la moralidad. Lo que nosotros podemos derivar ahora es que la valoración moral es un sentimiento placentero o doloroso que nos presenta de una cierta manera a una cierta personalidad o cualidades de ella. A continuación analizaré los elementos que parecen constituir a tal sentimiento.

Por una parte –como sabemos por la teoría de las pasiones de Hume– el sentimiento mismo no es el objeto de la valoración moral. Lo que valoramos por medio de ese sentimiento es el contenido del mismo, esto es, alguna percepción vívida. Además, necesitamos distinguir al sentimiento moral de otros sentimientos de placer o dolor. Sin esta distinción, el placer o desagrado morales tendrían a confundirse con los “juicios relativos a toda clase de gustos, sensaciones y belleza” (cf. T 3.1.2.3 SB 471) –observemos de paso cómo en este contexto Hume emplea como intercambiables “juicio” y “sentimiento”. La precisión que buscamos se enfrenta a la observación de que “bajo el término placer comprendemos sensaciones muy distintas unas de otras y que guardan entre sí únicamente la lejana semejanza necesaria para poder ser expresadas por el mismo término abstracto” (cf. T 3.1.2.4 SB 472). Así, son placenteros el vino o una composición musical e incluso podemos decir de cada uno que son *buenas* solo por el placer que nos producen, pero no lo son de la misma manera: «no diríamos que el vino es armonioso al oído ni que la música sabe bien al paladar».⁷⁸ Lo que marca la diferencia entre estas sensaciones

⁷⁸ Sigo los ejemplos de Hume (cf. T 3.1.2.4 SB 472), pero me parece que hoy en día podría decirse, literalmente, que un buen vino es armonioso o incluso que tiene buen ritmo. Respecto de la música tenemos formas no literales de

placenteras, según Hume, “es el modo diferente de sentir la satisfacción” o el desagrado. En consecuencia, tenemos que buscar la sensación distintiva de la moralidad, la cual constituye la sensación placentera de la virtud y, correspondientemente, la desagradable del vicio. Hume describe a la sensación moral en términos de *imparcialidad* y de su “sistema de las pasiones”. Cito extensamente el pasaje donde presenta a la idea de imparcialidad:

Las buenas cualidades del enemigo nos resultan nocivas, y pueden, sin embargo, seguir mereciendo nuestro aprecio y respeto. Sólo cuando un carácter es considerado en general y sin referencia a nuestro interés particular causa esa sensación o sentimiento en virtud del cual lo denominamos moralmente bueno o malo. Es verdad que los sentimientos debidos al interés y los debidos a la moral son susceptibles de confusión y que se convierten unos en otros. Así, nos resulta difícil no pensar que nuestro enemigo es vicioso, o distinguir entre su oposición a nuestros intereses y su real villanía o bajeza. Pero ello no impide que los sentimientos sean de suyo distintos: un hombre de buen sentido y juicio puede librarse de caer en esas ilusiones. De modo análogo, aunque sea verdad que una voz de timbre musical no es sino algo que proporciona naturalmente un determinado tipo de placer, nos resulta, sin embargo, difícil notar que la voz de un enemigo es agradable o admitir que es musical. En cambio, una persona de fino oído y con dominio de sí misma puede separar esos sentimientos y elogiar aquello que lo merece (T 3.1.2.4 SB 472).

El sentimiento moral se dirige exclusivamente a otra persona (la cual puede ser el propio agente), es decir, no tenemos sentimientos morales hacia entes no-humanos. Además, requiere que tal persona sea considerada “en general y sin referencia a nuestro interés particular”, esto es, de manera imparcial.⁷⁹ En efecto, según Alanen, lo que hace que una reacción afectiva se convierta en un sentimiento moral no es su contenido o el objeto de la percepción (un robo, un asesinato, salvar a alguien en peligro, etc.), sino el *modo* peculiar en que tal objeto es sentido, esto es, bajo un punto de vista general que no es relativo a los intereses del observador: “...el objeto particular que lo causa tiene que ser considerado en su aspecto general –no meramente según el modo cómo se relaciona con nuestros intereses” (Alanen, 196). Es difícil imaginar este punto de vista si se trata de algún suceso moralmente significativo cercano a nosotros, por ejemplo, que seamos víctimas de un robo violento o que seamos los beneficiarios de una acción heroica. Hume, en la

representarnos las emociones que nos produce apelando a sabores, etc. Creo que hay otras formas interesantes de ilustrar el punto de Hume: un auto puede ser “bueno” por su economía o potencia, un cuchillo por su filo o ergonomía, una silla por su originalidad o por su valor histórico, etc. ‘Bueno’, o cualquier otro término valorativo, tiene sentidos diversos que precisamos contextualmente. Hume, como veremos, distingue entre los usos de bueno por su carácter fenomenológico, *i.e.*, la forma distintiva en que experimentamos placer o dolor.

⁷⁹ En el pasaje citado de la IPM leemos, me parece, el mismo concepto de normativo de imparcialidad y generalidad que distingue al sentimiento moral, según Hume: “La noción de moral implica algún sentimiento común a toda la comunidad, que recomienda el mismo objeto a la aprobación general y hace que todos los hombres, o la mayoría de ellos, concuerden en la misma opinión o decisión sobre él. Implica también algún sentimiento tan universal y comprensivo como para abarcar a toda la humanidad y convertir las acciones y conductas, incluso de las personas más alejadas, en objeto de aplauso o censura según estén de acuerdo o en desacuerdo con esa regla de lo correcto que está establecida”.

cita anterior, nos indica que tal punto de vista equivale a una ponderación adecuada del carácter de la persona: ¿Estoy seguro que mi heroica salvadora actuó sin afán de fama o protagonismo? El punto de Hume es que el sentimiento moral requiere que suspendamos nuestros intereses particulares y ponderemos las cualidades propias o inherentes de quienes valoramos. Esta valoración moral requiere que *permitamos* que las cualidades de las personas nos afecten apropiadamente y que despierten así el placer o el desagrado correspondientes con ellas.⁸⁰ Esta idea es clave para sostener la interpretación del observador ideal.

Con el punto de vista general o imparcial Hume introduce el punto de vista de un sabio o de alguna autoridad reconocida como un estándar *normativo* relevante para explicar la motivación moral: “...nos resulta difícil no pensar que nuestro enemigo es vicioso, o distinguir entre su oposición a nuestros intereses y su real villanía o bajeza. Pero ello no impide que los sentimientos sean de suyo distintos: *un hombre de buen sentido y juicio* puede librarse de caer en esas ilusiones” (cf. T 3.1.2.4. SB 272, énfasis añadido). El punto de vista del «hombre de buen sentido o juicio» representa el estándar normativo para establecer el juicio general e imparcial sobre una persona.

Este estándar normativo tiene una función en el terreno de la moralidad análogo al de la experiencia en el terreno teórico. Como ha observado T. L. Beauchamp, Hume apela en varios pasajes al sabio (menciona T 1.3.13.12, 1.4.1.5 y 1.4.4.1; ver también el DNG) y dice que es un “modelo tácito de los estándares que debemos usar en el razonamiento y en la formación de creencias” (cf. Beauchamp 2008, 494-495). De acuerdo con Hume, tal estándar normativo consiste en un cierto ajuste de nuestras creencias con la experiencia, esto es, con la evidencia disponible para ellas: «el hombre sabio proporciona su creencia a la evidencia» (cf. EEH 10.1.4 SB 110). Esta idea parece recorrer el pensamiento de Hume y se refiere, según la expresión de Beauchamp, «a la experiencia como la categoría normativa fundacional», esto es, el único fundamento sólido de la ciencia y de todas las creencias empíricas (cf. Beauchamp 2008, 494 y T. Intr. 7). En el terreno teórico, la experiencia es la autoridad última para la formación de creencias y para la valoración de razonamientos –como hemos visto, su papel es especialmente

⁸⁰ Baillie caracteriza al sentimiento moral como una consideración “del carácter *per se* de la otra persona –lo universal más que a lo particular– tal como *ser un mentiroso*” (cf. Baillie, 140), pero esto impediría que pudiéramos juzgar a una acción particular, una mentira por ejemplo, como reprobable, aunque no consideremos en general a la persona como mentirosa. La noción de imparcialidad de Hume parece requerir solo que el sentimiento moral no surja de nuestros intereses particulares o que se haga desde una perspectiva general, no que en cada caso se juzgue de manera general a una persona (cf. Alanen 2006, 196; Rivera, 83).

relevante en relación con las creencias sobre causas y efectos. También es relevante para combatir prejuicios, creencias caprichosas, supersticiones a las que suelen ceder los agentes humanos por la facilidad con la que asienten a las determinaciones naturales de la asociación, hipótesis prejuiciosas, creencias de metafísica especulativa, etc. (cf. T 1.3.7.6 SB 101; Beauchamp 2008, 494 y Morris 91). En el terreno de la moralidad, la autoridad del sabio, como estándar normativo, consiste en una actitud semejante, pero dirigida hacia otros agentes y hacia sus acciones. En este caso, el estándar normativo promueve un punto de vista general o imparcial sobre otros agentes y sobre ciertas acciones, las que puedan tener algún valor moral por estar relacionadas con el agente que valora. Ésta es una moral que convierte a cada agente en un juez de los otros y a la vez lo hace objeto de los juicios de los demás. El estándar normativo que rige a los agentes morales es así construido y sostenido socialmente (cf. Baillie, 196-7).

El proceso por el cual los agentes adoptan dicho punto de vista general reúne elementos diversos y va progresando desde un punto de vista particular, hasta uno más general, el cual constituye finalmente el estándar normativo. En un primer momento, la contemplación o el interés por el carácter de otra persona, sus acciones o la persona misma producen en un agente un cierto placer o dolor, de acuerdo con la virtud o el vicio que reconoce en ellos. Estos sentimientos son particulares a cada agente y diversos entre sí, pero aquí interviene la simpatía, que lleva a los agentes a buscar y compartir puntos de vista que les son familiares respecto de lo que les produce placer o desagrado, esto es, respecto de lo que juzgan como vicio o como virtud (cf. Brand, 67; Taylor, 282). Esto es lo que explica Hume en el siguiente pasaje, sin mencionar a la simpatía: “Pero es fácil darse cuenta de que, siendo diferentes los placeres e intereses de cada persona en particular, es imposible que los hombres puedan convenir en sus sentimientos y juicios, a menos que elijan algún punto de vista común desde el que examinar su objeto y que pueda ser causa de que este último les parezca idéntico a todos ellos”. Este proceso lleva a los agentes a ajustar o calibrar sus puntos de vista de acuerdo con un estándar que resulta general y, por ello, “más constate y universal”. Una vez que se alcanza este punto de vista idéntico para todos, el estándar normativo queda instaurado y tiene el poder de superar a la fuerza de los intereses y placeres más particulares, aun cuando éstos sean más inmediatos. El placer y los intereses del punto de vista general afectan a los agentes con menor intensidad, pero son más persistentes, constantes y universales: “...aunque tales intereses y placer nos afecten con menor intensidad que los que nos son propios, como son, en cambio, más constantes y universales contrarrestan a éstos incluso en

la práctica y son los únicos criterios de virtud y moralidad admitidos en la especulación filosófica. Sólo ellos producen esa particular afección o sentimiento de que dependen las distinciones morales” (T 3.3.1.30 SB 591). El estándar normativo es, precisamente, ese *único criterio de virtud y moralidad*. En *la práctica*, como dice Hume, para cualquier agente real tales estándares serán parte de los dispositivos sociales de evaluación a su alcance, de modo que para un agente particular funcionan como criterios o normas contra los que calibra o ajusta gradualmente su sentido particular de lo virtuoso o de lo vicioso, de lo que es digno de ser alabado y lo que es digno de ser reprobado, con lo que eventualmente sus propias evaluaciones se vuelven constantes y estables (cf. Taylor, 282).⁸¹ Este punto de vista general, en efecto, consiste en el punto de vista del “observador ideal” que cada agente interioriza para hacer evaluaciones diversas sobre las acciones y pensamientos de otros agentes.

El *observador ideal*, según Hume, adopta la forma de un juez o, según estamos considerando, de la humanidad misma bajo la idea de la *universalidad* del criterio normativo. Este sabio, o autoridad normativa ideal, cuando juzga el carácter de otra persona lo hace con completa imparcialidad y en posesión de toda la información relevante. Me parece que el desarrollo más persuasivo de estas ideas lo elabora Hume en el ensayo “De la norma del gusto”.

Uno de los propósitos principales de Hume en DNG es rechazar la “clase de filosofía que cercena toda esperanza de éxito en tal intento [el de encontrar una *norma del gusto*, i.e., una regla que permita conciliar los diversos sentimientos de la gente, o al menos una decisión que confirme un sentimiento y condene otro], y que representa la imposibilidad de llegar a ninguna norma del gusto” (cf. DNG §6-7, 221-222).⁸² Hume entiende en dos sentidos la noción de norma. En el primero se refiere a un tipo de regla para igualar los sentimientos de las personas y con el segundo a un criterio o estándar para separar de algún modo a los sentimientos aceptables de los inaceptables. Este último sentido presupone que en ciertas circunstancias algunos sentimientos serán correctos y otros no, mientras que el primer sentido parece más orientado a mostrar qué elementos tendrían que convergir para que en ciertas circunstancias los agentes experimentarían el mismo sentimiento. Ambos sentidos son normativos, pero su función es diferente. Hume se

⁸¹ En el terreno de las normas estéticas el procedimiento es similar, tal como lo reconoce Jones, y constituyen la fuente de los estándares estéticos que son socialmente útiles para las evaluaciones respectivas (Jones, 263).

⁸² Dada la diversidad de ediciones de este ensayo y que, hasta donde sé, no existe una edición canónica, en mis referencias indico el número de párrafo y la página de la edición en español que cito.

propone rechazar una posición en contra de estos sentidos de la norma del gusto. Ahora bien, al defender estos sentidos de la norma del gusto equipara al sentimiento con el juicio, por medio de la norma del gusto que se propone defender. La posición que rechaza confina al sentimiento a la esfera irreductible de la vida mental de cada agente, donde no hay más norma que los propios sentimientos:

Es muy grande, se dice, la diferencia que separa el juicio del sentimiento. Todo sentimiento es correcto, porque el sentimiento no se refiere a nada más allá de sí mismo, y es real siempre que una persona sea consciente de él. En cambio, no son correctas todas las determinaciones del entendimiento, porque se refieren a algo que está más allá de ellas mismas, al ingenio, a los hechos reales, y no pueden siempre conformarse a esa norma (DNG §7, 222).

Así, la posición que Hume rechaza presupone que hay una brecha entre el sentimiento y los juicios, que es la misma brecha que habría entre los estados cognitivos y los no-cognitivos de la mente. Mientras que el sentimiento, para usar la caracterización del *Tratado*, es una «existencia originaria que no representa nada», el juicio sí representa algo y en la medida en que la representación es adecuada o correcta, el juicio es correcto, pero en la medida en que no lo es, el juicio es incorrecto. Bajo esta posición, la apreciación de la belleza sería meramente una cualidad de la mente la cual, en la medida en que es pensada o percibida por la consciencia, es «correcta», o por lo menos no puede ser incorrecta:

La belleza no es una cualidad de las cosas en sí, y cada mente percibe una belleza diferente. Puede incluso ocurrir que una persona perciba como deformidad algo que para otra es belleza, y cada individuo tiene que aceptar su propio sentimiento sin pretender regular el de todos los demás (DNG §7, 222).

El sentimiento de belleza, en consecuencia, no está ni puede ser regulado, es decir, ninguna norma o regla puede aplicársele porque no habría ningún parámetro externo a la mente que la percibe para decidir si tal o cual apreciación de belleza es correcta o si alguien falla en no experimentarla. Ésta es la posición contra la que Hume se propone argumentar.

Según entiendo el argumento de Hume, su posición no es inconsistente ni opuesta a la posición que defiende tanto en el *Tratado* como en la IPM. Para él, los sentimientos son susceptibles de ser influenciados de manera consistente por factores sociales, incluso si los agentes no son del todo conscientes de tal influencia (v. infra §5.2) Entre estos factores están normas y reglas morales, y su efecto sobre la voluntad de los agentes es motivar conductas y pensamientos consistentes entre sí y con dichas reglas. Ésta es la posición que defiende también en el ensayo DNG y me parece que la relevancia de este texto es la claridad con la que aborda este tema y la explicación de cómo interiorizamos las normas por medio de la experiencia y la

educación. Pero Hume también reconoce que hay una cierta determinación de nuestra naturaleza a buscar tales estándares, por ejemplo cuando afirma que «es natural que busquemos una norma del gusto», precisamente en el doble sentido que indiqué arriba (cf. DNG §6, 221) o acusa a quienes sostienen la posición opuesta a la suya «de olvidar por completo el principio de igualdad natural de los gustos» (cf. DNG §8, 223). La primera idea se refiere a una cierta tendencia natural que hay entre los seres humanos para acordar o seguir pautas de pensamiento y conducta similares. Creo que podemos aceptar con pocas reservas esta idea. La segunda se refiere a que la disposición común de nuestros «órganos» y, en general, a la sospecha bien fundada de que los seres humanos estamos *cableados* del mismo modo y respondemos de maneras similares a estímulos similares. Como hemos visto, las diferencias entre los seres humanos se explican por las diferencias culturales y éstas por diversas contingencias. No obstante, en DNG va a argumentar que además de estas determinaciones naturales podemos reconocer otras derivadas de la experiencia y la educación.

Hume afirma al respecto que las reglas de composición «tienen por fundamento el mismo que el de todas las ciencias prácticas: la experiencia, no son nada más que observaciones generales sobre lo que universalmente se ha hallado que complace en todos los países y en todas las épocas» (DNG §9, 223). A favor de esta idea menciona a las preferencias por ciertas poesías y poetas que reconoce en todos los pueblos, como Homero, Terencio o Virgilio, quienes “imperan, universal e indiscutiblemente, sobre las mentes de los hombres” (cf. DNG §26, 232).⁸³ Pero no es suficiente con contar con ejemplos como éstos porque muchos factores pueden afectar a la percepción. Podría suceder que padecemos de una mala disposición para percibir ciertas cualidades de los objetos relativas a la belleza por no tener «serenidad de la mente, una actitud de concentración y la debida atención a un objeto (cf. DNG §10, 224). Ciertas disposiciones de carácter, como la envidia o los celos podrían alterar nuestra sensibilidad estética hacia un artista rival (cf. DNG §11, 224) o quizá no contamos en alguna circunstancia con las condiciones adecuadas para percibir las cualidades de un objeto (cf. DNG §13, 225). Aún así, la «causa» más común que Hume encuentra «por la cual muchas personas no experimentan el adecuado sentimiento de belleza es por la falta

⁸³ Hume contrasta en este pasaje a estos poetas con las obras de los filósofos, Aristóteles, Platón, Epicuro y Descartes, que pueden «sucesivamente dar paso una a otra», es decir, que la filosofía no es universal como la poesía. Todavía más, dice en seguida que la “filosofía abstracta de Cicerón ha perdido su crédito, mientras que la vehemencia de su oratoria sigue despertando nuestra admiración”. Veremos más adelante que, en efecto, Hume considera que los juicios del entendimiento son más cambiantes que los uniformes juicios del sentimiento.

de esa *delicadeza* de la imaginación que se requiere para proporcionar una sensibilidad para esas emociones más refinadas» (cf. DNG §14, 225). De modo que si se corrigen o no se padece alguno de los inconvenientes mencionados para experimentar adecuadamente el sentimiento de belleza, entonces la imaginación es la aliada de la experiencia que puede despertarlo. En efecto, Hume afirma que el estándar del gusto «no lo fija en ningún caso el razonamiento *a priori* ni es una conclusión abstracta del entendimiento, suponiendo que existan principios eternos e inmutables para derivar de ellos la norma de belleza» (cf. DNG §9, 223). Esta afirmación es consistente con su pensamiento porque, como hemos visto en el caso del círculo citado antes, no hay ninguna cualidad del círculo de la que podamos derivar la belleza. Ni la proporción o armonía le son propias, pues parece que estas cualidades son también cualidades en la mente de los agentes y que, en consecuencia, pueden ser explicadas por el argumento que estamos considerando. Por razones semejantes, la belleza no es reducible a cualidades sensoriales, porque entonces la percibiríamos empleando la vista o el oído, etc. Así que, por descarte, la percepción de la belleza tiene que provenir de la imaginación, que resulta así la capacidad de la mente adecuada para establecer a las normas estéticas y morales. La imaginación, con todo, requiere ser dirigida o entrenada, como sucede comúnmente en las ciencias prácticas.

Así, hablando de «las bellezas del arte de escribir» y cómo éstas por alguna metodología se han “reducido a principios generales... modelos excelentes”, las cuales han eliminado «los diferentes grados del gusto», Hume deriva la idea de que si dichos principios y modelos son compartidos con alguien empeinado en su propio criterio adverso a ellos eventualmente llegará a apreciar “la delicadeza que se requiere para apreciarlos”:

...cuando le mostramos un principio del arte aceptado; cuando ilustramos este principio con ejemplos que, desde su propio gusto particular, considera conformes al principio; cuando demostramos que el mismo principio puede aplicarse al presente caso, aunque él no perciba o sienta su influencia, tendrá que concluir, en conjunto, que la falta es suya, y que carece de la delicadeza que se requiere para ser sensible a toda belleza y a todo defecto en una composición o discurso (DNG §16, 227).

Con esta idea Hume introduce la idea de comunidad y experiencia que son depositarios de ciertos criterios, los cuales transmite a un no iniciado en ellos por medio de ejemplos, demostraciones y contrastaciones. El iniciado en esta comunidad realiza gradualmente un proceso de internalización de los principios y modelos reputados como la norma de belleza, de modo que con el tiempo llega a ejercerlos y percibirlos. Una persona entrenada de esta manera no puede sino aceptar y aplicar tales normas o, me parece, ser proscrita de la comunidad que las tiene y cultiva. Lo que la persona

hace es interiorizar cierta forma de percibir una norma relativa a ciertos objetos, como un poema, una pintura, una acción humana, etc. Con el tiempo, para esta persona la percepción de tales normas será indistinta de la percepción de ciertas cualidades en el rango apropiado de objetos (textos, por ejemplo, o acciones humanas de cierto tipo) y posiblemente no pueda ya distinguir entre las cualidades que percibe sensiblemente y la norma evaluativa. Pero esta percepción no es otra cosa que un acto de la imaginación entrenada apropiadamente –y no, por ejemplo, un acto de intuición de alguna facultad mental indeterminada. Hume nos explica en los siguientes términos este proceso de educación de la imaginación:

...aunque naturalmente exista una gran diferencia respecto a la delicadeza entre una persona y otra, nada tiende a seguir incrementando y mejorando este talento como la práctica de un arte determinado y el frecuente estudio o contemplación de una determinada clase de belleza. Cuando por primera vez se presentan objetos de cualquier clase ante el ojo o la imaginación, el sentimiento que los acompaña es oscuro y confuso, y la mente es incapaz en gran medida de pronunciarse respecto a sus méritos o defectos. El gusto no puede percibir las diversas excelencias de su realización, y mucho menos distinguir el particular carácter de cada excelencia y verificar su calidad y grado. [...] Pero, si se permite que esa persona [sin práctica] adquiera experiencia en tales objetos, su sensibilidad se hace más precisa y refinada. No sólo percibirá las bellezas o defectos de cada una de sus partes, sino que señalará lo que distingue cada una de sus cualidades, y le dedicará la alabanza o el reproche adecuados (DNG §18, 227-228).

Este proceso de educación de la imaginación requiere, en efecto, de una comunidad que indique las reglas y los modelos, de práctica constante por parte de los educandos y de una cierta actitud para aprehender aquel lenguaje o criterios que la comunidad emplea. En la IPM Hume nos explica cómo el uso del lenguaje juega un papel decisivo para expresar pensamientos y aplicar los estándares normativos: “El lenguaje se adapta a los sentimientos generales. Ciertos términos generales de censura y aprobación son adecuados para expresarlos. Esto propicia que ciertas acciones se realicen de acuerdo con ellos” (IPM 9 SB 273-274).⁸⁴ En conjunto, estas ideas le permiten a Hume sostener que el sentimiento de belleza es comparable con la capacidad de juzgar. En otras palabras, que experimentar la belleza de un objeto equivale a ser capaz de juzgarla como bella, de manera correcta y en consonancia con los cánones compartidos por una cierta comunidad. Hume sostiene esto equiparando a la producción artística con el juicio respectivo: “En resumen: la misma capacidad y destreza que la práctica proporciona para la ejecución de una obra se requieren también, por el mismo medio, para juzgarla” (DNG §18, 228). La práctica sostenida y, al parecer, recibida favorablemente por la comunidad correcta, conduce a un cierto

⁸⁴ Hume parece decir que el propio lenguaje moral es una expresión de la convergencia de los sentimientos morales de acuerdo con las normas de aprobación o desaprobación: “En el lenguaje moral se expresan sentimientos comunes a toda la humanidad. Quien habla así se aleja de su situación privada y particular” (v. EPM, 9 SB 272-273).

refinamiento que *no* puede evitar juzgar el objeto de su experiencia. Esta idea hace más fuerte la tesis de Hume con la que compara a la norma del gusto con un juicio del entendimiento: “Rara vez ocurre, si es que ocurre alguna vez, que una persona sensata que tenga experiencia en cualquier arte no sea capaz de juzgar su belleza, y no es menor raro encontrar a alguien que tiene buen gusto sin un sano entendimiento” (DNG §22, 230-231). Además, los obstáculos que enfrenta el entendimiento para juzgar son los mismos que sortea el sentimiento para apreciar la belleza: “Es bien sabido que, en toda cuestión sometida al entendimiento, los prejuicios destruyen el buen juicio y pervierten todo el funcionamiento de las facultades intelectuales. No son éstos menos contrarios al buen gusto, ni influyen menos en corromper nuestro sentimiento de la belleza” (DNG §22, 230). Hume no duda que este proceso educativo afina y prepara la disposición natural y universal de los agentes humanos para percibir la belleza. Las diferencias al respecto provienen de una actitud negligente o a factores contingentes en la educación y entrenamiento de la imaginación. Ahora bien, a pesar de su universalidad de las disposiciones humanas y de la capacidad para apreciar ciertas cualidades sensibles correlativas con la norma del gusto, «la verdadera norma de la belleza» es enteramente una cualidad mental y, al mismo tiempo, es compartida consistentemente por una cierta comunidad:

Una sólida sensatez, unida a la delicadeza de sentimientos, mejoradas por la práctica, perfeccionadas por la comparación y liberadas de todo prejuicio, son lo único que puede proporcionar a alguien estas valiosas cualidades, y el veredicto conjunto de tales personas, cuando se las encuentra, constituyen el verdadero canon del gusto y la belleza (DNG §23, 231)

La misma analogía que hemos visto antes entre el sentimiento de belleza y el sentimiento moral la sigue sosteniendo Hume en DNG en términos semejantes, es decir, afirma que también dan lugar a juicios.

No obstante, Hume no desarrolla cuáles son los elementos que pueden compararse entre el sentimiento moral y el sentimiento de la belleza, pues asume que son el producto del mismo tipo de operación mental y, por lo tanto, que son el mismo tipo de juicio. De manera que parece concentrar su atención en el hecho de que son susceptibles de los mismos procesos de corrección cognitiva que los juicios del entendimiento, según explicaré en la siguiente sección. Antes de ello, elaboraré los elementos de la analogía. Según estamos considerando, aun cuando el sentimiento moral es una cualidad mental que, en tanto que juicio, se *refiere* a una norma moral, como la justicia, la virtud, la belleza de una acción, etc., requiere de la percepción de ciertos objetos y propiedades sensibles. Como en el caso de la belleza, los agentes cuentan con una cierta

experiencia acumulada que orienta su percepción. Esta experiencia, según hemos visto, ya ha sistematizado las reglas o principios y destacado como modelos ciertas acciones, ciertos caracteres, etc., de manera que los agentes interiorizan lo relativo a estas entidades. Las normas morales son derivadas de la percepción de estas entidades por medio de la imaginación que, de este modo, contribuye a la formación del sentimiento o juicio moral. Las entidades morales a las que se refieren los juicios morales no se reducen a las entidades percibidas sensiblemente, no equivalen a ellas, porque son derivadas por la imaginación y, por lo tanto, son un producto de ella. David Wiggins nos explica que Hume nos ofrece así una visión naturalista de las entidades morales que «no necesariamente genera una explicación de ‘bueno’ del tipo que Moore anatematizó por cometer ‘la falacia naturalista’, la falacia de asimilar sujetos de otro asunto a los asuntos de la ciencia experimental. En la explicación humeana, la extensión de ‘bueno’ y otros predicados de valor no están determinados del modo propio de una ciencia natural» (cf. Wiggins 2006, 34 nota 2). Es decir, para Hume las entidades morales, no pueden ser reproducidas por las metodologías características de las ciencias naturales, sin embargo, estas entidades están determinadas por la naturaleza humana y las condiciones en que percibe y produce conocimiento, lo que da lugar a respuestas universales y consistentes entre todos los agentes, como las reglas de la justicia. Apreciamos la universalidad y consistencia en la aplicación de las reglas morales en el terreno práctico, donde los juicios morales contribuyen a distinguir y caracterizar actos como injustos, viciosos, buenos, etc. En la sección 6.2 defenderé que estas entidades no son menos “raras”, siguiendo la objeción que plantea Mackie (Mackie 2000) con su teoría del error, que, digamos, las relaciones causales según las propias ideas de Hume.

Una de las afirmaciones más interesantes del ensayo DNG es que sostiene que los juicios morales son más constantes o estables que los juicios del entendimiento: “No son lo mismo los principios morales que las opiniones especulativas de cualquier tipo. Éstas (*sic*) últimas están en cambio y revolución continuos. El hijo se adhiere a un sistema distinto que el padre. Es más, apenas hay alguien que pueda presumir de gran constancia y uniformidad a este respecto” (DNG §33, 235). En seguida Hume menciona la cantidad de formas de divergencias y errores que encontramos en el pensamiento especulativo, producto del entendimiento. No obstante estos errores, reconoce que solemos de todos modos reconocer el mérito estético de las composiciones especulativas, lo cual fortalece su afirmación a favor de la estabilidad y universalidad de los

principios estéticos. En cambio, en la esfera del sentimiento moral, y por extensión, de su expresión como juicio, encuentra más constancia y regularidad:

En cambio, se requiere un violento esfuerzo para cambiar nuestro juicio sobre las costumbres, y para suscitar sentimientos de aprobación o desaprobación, de amor o de odio, diferentes de aquellos con los que la mente ha estado familiarizada por prolongado hábito. Y, cuando tenemos confianza en la rectitud de esa norma moral, por la que se guía nuestro juicio, somos justamente celosos de la misma, y no pervertiremos los sentimientos de nuestro corazón para caer en la complacencia con ningún autor, sea el que fuere (DNG §33, 235)

La uniformidad del sentimiento y juicio moral se debe, en parte, al hábito y la costumbre, esto es, a una cierta actitud hacia el pensamiento y la acción que se ha consolidado con el tiempo y el reconocimiento sociales. Sin embargo, Hume también parece describir una actitud poco racional hacia tales sentimientos y juicios. Su posición, por ejemplo, parece dogmática respecto de no aceptar ninguna opinión contraria, «sea de autor que sea». Quizá está manifestando una actitud más cautelosa que algunos pasajes del *Tratado* o del EEH donde descalifica sin demasiados miramientos a la superstición y el pensamiento religioso, porque en el siguiente párrafo nos dice que de “todos los errores especulativos, los que conciernen a la religión son los más excusables en las composiciones de talento” (cf. DNG §34, 235).⁸⁵ Creo que Hume pudo ofrecer una razón más interesante para esta uniformidad. Se puede explicar esta constancia y regularidad del sentimiento moral porque *expresa* a la naturaleza humana, la cual, como hemos visto, es universal. Esta idea recorre el pensamiento de Hume y complementa consistentemente lo que señala aquí sobre la constancia del sentimiento moral. Además, como desarrollaré en la siguiente sección, Hume también reconoce que el sentimiento estético –y por extensión- el moral, es objeto del mismo tipo de revisión racional que aceptamos sin reservas para los juicios del entendimiento. Consideremos ahora el segundo rasgo distintivo del sentimiento moral, anunciado antes, *el sistema de las pasiones*.

En efecto, después de la imparcialidad, Hume dice que el segundo rasgo distintivo del sentimiento moral lo aporta su “sistema de las pasiones”:

En *segundo* lugar podemos recordar ahora el precedente sistema de las pasiones, con el fin de poner de relieve una diferencia aún más notable entre nuestros dolores y placeres. Las pasiones de orgullo y humildad, de amor y de odio, se producen cuando se nos presenta algo que guarda relación con el objeto de la pasión y que a la vez produce por separado una sensación de la pasión. Pero la virtud y el vicio están acompañados por estas circunstancias: tienen que estar necesariamente situados o en nosotros mismos o en otros, así como excitar placer o desagrado; por tanto, deben originar una de las cuatro pasiones antes mencionadas. Lo que distingue a la virtud y el vicio del placer o dolor

⁸⁵ Un ejemplo de la temeraria opinión opuesta, proveniente del *Tratado*: “Hablando en general, los errores en materia de religión son peligrosos; los de la filosofía solamente ridículos” (T 1.4.7.13 SB 272).

procedente de objetos inanimados consiste en que, a menudo, no están relacionados con nosotros. Este es, posiblemente, el efecto más notable que la virtud y el vicio tienen sobre la mente humana (T 3.1.2.5 SB 473).

De acuerdo con este argumento, el sentimiento moral tiene la misma estructura psicológica de la doble relación que caracteriza a las pasiones indirectas: la idea de uno mismo y la idea de otra persona considerada de una cierta manera. Ésta es una relación que no está siempre presente cuando se experimenta placer o dolor.

Solemos experimentar orgullo o humildad por objetos inanimados como una casa o alguna otra posesión y, en tanto meros sentimientos placenteros o desagradables, carecen de valor moral. En efecto, el orgullo o placer que tengo por poseer la casa más bonita de mi calle carece de importancia moral y, desde el punto de vista de la teoría de las pasiones de Hume, es simplemente un sentimiento placentero relativo a una posesión mía. Correspondientemente, la humildad o desagrado que me podría producir poseer la casa más fea o pobre de la calle carecería también de valor moral y su manifestación es solo la manifestación de una sensación de desagrado. Respecto del amor o el odio podemos hacer consideraciones similares, cuando se refieren a objetos inanimados. El placer de amar un lugar para tomar café, o el desagrado de odiar una cierta pieza musical carecen de valor moral alguno y no son más que ciertas manifestaciones de la psicología personal de cada agente. Pero si la pasión se refiere a una persona, entonces su naturaleza es otra y, de acuerdo con la primera característica de los sentimientos morales, podría ser de calidad moral. Hume considera que el efecto más importante del vicio y de la virtud podría ser que hacen de otra persona el objeto de consideraciones especiales, esto es, morales. Este proceso de atención moral requiere, según su teoría, de que ciertas pasiones indirectas se despierten relacionando a ciertos agentes bajo una cierta perspectiva emocional, según se aprecia en el siguiente pasaje:

Las pasiones de orgullo y humildad, de amor y odio, se producen cuando se nos presenta algo que guarda relación con el objeto de la pasión y a la vez produce por separado una sensación de la pasión. Pero la virtud y el vicio están acompañados por estas circunstancias: tienen que estar necesariamente situados o en nosotros mismos o en otros, así como excitar placer o desagrado; por tanto, deben originar una de las cuatro pasiones antes mencionadas. Lo que distingue a la virtud y el vicio del placer o dolor procedente de objetos inanimados consiste en que, a menudo, no están relacionados con nosotros. Éste es, posiblemente, el efecto más notable que la virtud y el vicio tienen sobre la mente humana (T 3.1.2.5 SB 473).

Con todo, lo esencial al sentimiento moral es la actitud de imparcialidad que adoptamos sobre otros agentes. Por ejemplo, amar a un amigo podría carecer de valor moral y aunque el grado o la intensidad del sentimiento podría ser muy distinto a amar un lugar para tomar café, no se trata de un sentimiento moral a menos que el sentimiento reúna también el juicio general de imparcialidad, por ejemplo, amarlo a pesar de que conocer sus fallas, malas disposiciones, etc., no disminuyan nuestro afecto por él. Asimismo, el odio podría tener valor moral si se dirige a una persona a quien tal sentimiento no nos hace escatimarle sus méritos y buenas cualidades. La humildad y el orgullo por otras personas, de manera similar, requieren del juicio general de imparcialidad para que adquieran el carácter de pasiones morales. Recordemos cómo el punto de vista general o imparcial dirige aún la consideración sobre un enemigo: "...nos resulta difícil no pensar que nuestro enemigo es vicioso, o distinguir entre su oposición a nuestros intereses y su real villanía o bajeza" (cf. T 3.1.2.4. SB 272). El juicio general de imparcialidad, en todos estos casos, parece funcionar así: aporta un punto de vista impersonal de manera que hipotéticamente cualquier persona que ocupe ese punto de vista, tengo o no relación con la persona a la que se dirige el juicio, aceptaría sin reservas el sentido y el valor del juicio en cuestión. Vería que el juicio general de amor, odio, humildad u orgullo por la persona puede ser recibido sin objeciones o reservas. Así, las pasiones morales y los motivos morales subsecuentes, en tanto que son pasiones indirectas, están estructuradas bajo un juicio general que dirige a la pasión y al motivo. Para resaltar los elementos de esta estructura, consideremos los siguientes contraejemplos, en los que falta el juicio general: Nuestro amor por nuestro amigo es tal que nos hace pasar por alto rasgos indeseables de su personalidad, los cuales tratamos de presentar como idiosincrasias encantadoras o minimizamos su efecto y ciegos a ellos promovemos sus intereses, a pesar de que tales rasgos indeseables podrían ser dañinos para otras personas. En este caso, el punto de vista general e imparcial reprobaría la omisión de tales rasgos y podría cuestionar u objetar que por amor promovamos sus intereses –tal vez no cuestionaría nuestro amor por el amigo, pero dudaría que se trate de un amor con valor moral. Similarmente, si nuestro orgullo por nuestro político local pasa por alto su actuar fraudulento o negligente, bajo el criterio del juicio general e imparcial podríamos objetar los motivos del orgullo y cuestionaríamos, digamos, el motivo para homenajearlo o promoverlo. Así, podemos aceptar que la pasión moral y el motivo subsecuente requieren de una consideración especial, el punto de vista general e imparcial sobre otra persona,

para distinguirlos de pasiones y motivos no morales. Hasta aquí el argumento por el que Hume caracteriza al sentimiento moral.

Hasta donde puedo darme cuenta, este argumento no ha recibido la apropiada atención por parte de los estudiosos de Hume, pero es importante porque con él enlaza su teoría de las pasiones y su teoría de la moralidad. Él mismo lo asienta explícitamente unas secciones más adelante en el *Tratado*: “Ninguna acción puede ser moralmente buena o mala, a menos que una pasión natural o motivos nos impelan a ella, o nos desvíen de ella; y es evidente que la moralidad tiene que ser susceptible de todas las mismas variaciones que son naturales a las pasiones” (T 3.2.6.9 SB 532). Pero sobre todo este argumento es importante porque con él sostiene el sentimiento moral se relaciona con alguna de los dos pares de pasiones básicas (orgullo y humildad, amor y odio). La estructura de estas pasiones asegura que el sentimiento se refiera o bien a uno mismo o bien a otra persona. Para un agente humeano, en consecuencia, no hay lugar para consideraciones abstractas sobre la moralidad. Los juicios morales, que son expresados mediante estos sentimientos valorativos, se refieren *siempre* a algún agente.

5. El naturalismo moral de Hume: la agencia y la moral

En este capítulo examino cómo el naturalismo de Hume explica las prácticas morales sociales con base en las determinaciones naturales estudiadas en los capítulos anteriores, a saber, las relativas a la formación de creencias, las pasiones, los motivos para actuar y los juicios morales. Particularmente, estudiaré dos problemas centrales. El primero consiste en reconocer que las determinaciones morales naturales son insuficientes para regular la vida social, de modo que se requiere de recursos no naturales, esto es, *artificiales* según sus términos, para complementar o reforzar a los primeros. El segundo problema consiste en cómo los agentes conforman su conducta a las reglas. Con el primer problema me refiero a los límites de las *virtudes naturales* para explicar las regulaciones sociales. Ante esto, Hume argumenta que se requieren de recursos creados socialmente para reforzar a los primeros y explicar las regulaciones sociales: “...nuestro sentimiento de la virtud no es natural en todos los casos, sino que existen algunas virtudes que producen placer o aprobación gracias a un artificio o proyecto debido a las circunstancias y necesidades de los hombres” (T 3.2.1.1 SB 477). Tales recursos comprenden a las reglas de la justicia, las promesas y el respeto a la propiedad privada. Con todo, Hume observa muy agudamente que tales recursos requieren ya de la vida social, por ejemplo, que los agentes los reconozcan como reglas que limitan su conducta, pero esto sólo es posible si las propias reglas ya están en funcionamiento. Para responder esta circularidad, Hume propone una de sus teorías más interesantes y, me parece a mí, vigentes: la justicia, las promesas y el respeto a la propiedad privada no surgen por acuerdos explícitos, reflexivos, sino por acuerdos tácitos, por actos de convención celebrados sin el auxilio del lenguaje. De este modo, Hume coloca a tales acuerdos en un espacio hipotético previo al surgimiento de las sociedades humanas, al tiempo que los convierte en las condiciones que las hacen posibles. Los agentes siguen tales acuerdos tácitos por la costumbre y el hábito asociados con el placer que obtienen de ello, y no por alguna consideración racional al respecto. De este modo, el escepticismo de Hume sobre la razón también la excluye de participar en las convenciones que dieron origen a las reglas morales básicas. Tales reglas son también expresión del sentimiento moral.

El tratamiento del segundo problema, cómo los agentes conforman su conducta a las reglas, me permitirá relacionar aspectos relativos a la psicología individual con las reglas de la justicia y de las promesas, pero considerando criterios normativos generales. Exploraré aquí cómo

las determinaciones morales naturales se relacionan con las reglas sociales. La filosofía de la moralidad de Hume, por contraste con la kantiana, concibe a los agentes morales preocupados constantemente por el punto de vista de los otros. Esta preocupación tiene un importante papel regulativo, porque los dispositivos morales de control, los criterios normativos más generales mencionados, se refieren a la vergüenza, la alabanza y la fama. Los agentes humeanos no están preocupados por el deber, o lo están sólo en la medida en que a su vez está involucrada la vergüenza, la alabanza o la fama, que serían así los dispositivos morales que regulan en último término su conducta. Específicamente, estos agentes respetan las reglas de la justicia, las promesas y de la propiedad privada, desde un punto de vista moral, pensando y actuando de acuerdo con los dispositivos morales mencionados. Creo que estas ideas se relacionan bien con la teoría del observador ideal presentada en el capítulo anterior.

En efecto, la constante atención y preocupación por el punto de vista de los otros da lugar en un agente individual a una especie de plataforma psicológica donde se asienta el observador ideal. Éste no necesariamente es impersonal, pero sí es una figura que juzga al agente desde su interior, según los parámetros sociales mencionados. De este modo, el agente humeano interioriza los dispositivos morales mencionados.

5.1. La moral y la psicología individual

El naturalismo filosófico de Hume descansa en la siguiente observación, muy segura y contundente:

Es evidente que todas las ciencias se relacionan en mayor o menor grado con la naturaleza humana, y que aunque algunas parezcan desenvolverse a gran distancia de ésta regresan finalmente a ella por una u otra vía. Incluso las matemáticas, la filosofía natural y la religión natural dependen de algún modo de la ciencia del hombre, pues están bajo la comprensión de los hombres y son juzgadas según las capacidades y facultades de éstos (T I 4 SB xv).⁸⁶

Hume fundamenta aquí un *proyecto* filosófico que depende, en general, de cómo los seres humanos, «los hombres», obtienen conocimiento y cómo lo evalúan. De acuerdo con Stroud, este proyecto fue «revolucionario» para la agenda de la investigación filosófica (cf. Stroud p. 23). El principal recurso de esta revolución fue la teoría de las ideas que, como hemos visto, suministra los elementos básicos para explicar la formación de las creencias, las pasiones y el sentimiento

⁸⁶ Sigo la referencia de la edición de David Fate Norton y Mary J. Norton. La edición en español de Félix Duque y otras referencias de este pasaje indican p. XX.

moral. Además, también implica un cierto modelo de la mente. Stroud explica que con base en estos antecedentes Hume:

...se concentra en indagar qué es lo que debe añadirse, cuáles son las verdaderas características de los seres humanos que pueden explicar por qué piensan, sienten, y actúan como lo hacen. [...] Intenta hacerlo mediante una investigación experimental sobre el pensamiento, el sentimiento y el comportamiento humanos, y no mediante lo que considera un teorizar tradicional y a priori acerca de la “naturaleza humana” o la “esencia del hombre” (Stroud p. 23)

Según Stroud, el proyecto de Hume es un proyecto descriptivo de la naturaleza humana y de las condiciones en que se producen diversos pensamientos, sentimientos y acciones. Además, es un proyecto anti-metafísico en el sentido en que no presupone nada *a priori* acerca de su objeto de estudio, la naturaleza humana. Quizá esta revolución se aprecia mejor enfatizando este último rasgo: Hume no presupone nada sobre la naturaleza humana que no provenga directamente de la «experiencia y la observación». Hume confía en que el conocimiento que se obtenga de ese modo podrá ampliarse a “todas las ciencias que más de cerca conciernen a la vida del hombre” (T I 6 SB xvi), las cuales incluirían lo que hoy identificaríamos como epistemología, psicología, ética, antropología y aun sociología. Stroud también nos dice que la investigación de Hume pretende *complementar* a su teoría de las ideas. En efecto, esta primera teoría nos ofrece los materiales básicos con que operan las diversas facultades de la mente, las impresiones y las ideas, de cuya combinación surgen todas las percepciones, esto es, “cualquier cosa que pueda estar presente ante la mente, así sea que empleemos nuestros sentidos, o sean accionados con las pasiones, o ejercidos por medio del pensamiento y la reflexión” (T A 5; SB 647).⁸⁷ En consecuencia, Hume se propone explicar, con base en unos pocos elementos, el universo de pensamientos humanos. Ampliando sus ideas relativas a las valoraciones morales, yo me refiero a este esfuerzo explicativo con su expresión «la flexibilidad de los sentimientos», porque me parece que la tarea más importante de su investigación es describir cómo los agentes humanos hacen valoraciones y cómo establecen normas, tanto para conocer al mundo como para regular sus acciones. En consecuencia, siguiendo la presentación de Stroud, yo diría que su proyecto naturalista consiste en explicar, con base en los elementos básicos de su teoría de las ideas, cómo los agentes humanos hacen valoraciones y establecen normas diversas. Con base en estas ideas, en este capítulo extenderé las ideas de Hume sobre las determinaciones naturales de la motivación hasta la esfera de la moral (el conjunto de

⁸⁷ Esta referencia corresponde a “An Abstract of a Book lately Published, entitled *A Treatise of Human Nature, &c.*”, escrito por el propio Hume y publicado anónimamente en 1740, el cual no está incluido en la edición preparada por Félix Duque. La traducción es mía.

las reglas y principios morales mediante el cual los miembros de una sociedad regulan su conducta de acuerdo con ellas).

Según la caracterización que he defendido en la sección anterior, los sentimientos morales tienen la estructura de una pasión indirecta. Por ello, están también constituidos por elementos cognitivos y sobre ellos descansan los juicios morales. En efecto, adoptar un punto de vista general e imparcial requiere que los agentes adopten una cierta actitud hacia otros agentes: que por lo menos «no consideren sus propios intereses». Así, la única diferencia entre cómo se experimenta ordinariamente una pasión indirecta y los sentimientos morales parece ser el punto de vista imparcial y general de estos últimos. L. Alanen marca la misma distinción cuando observa que “[L]a distinción entre las pasiones y los juicios morales en Hume es muy delgada y depende, básicamente, en sus objetos y en el punto de vista desde el que son considerados” (Alanen, 194). Con todo, como hemos visto en el caso de las creencias y de las pasiones, también en el caso del sentimiento moral sus elementos cognitivos no son suficientes para que sea objeto de la razón. Estos elementos cognitivos, ciertas *ideas* que nos informan de algún aspecto o algunos aspectos de un objeto o un estado de cosas, están presentes en prácticamente todos los estados mentales. Lo que hace que el estado mental correspondiente a las pasiones o al sentimiento moral sea o no objeto de la razón es de qué manera la razón contribuye a su formación. Como es característico en el pensamiento de Hume, lo que marca la diferencia entre estados mentales es su origen o las causas de las que surge. En el caso de las pasiones (y aún de las creencias) la razón no es decisiva para que surja el estado mental correspondiente. El sentimiento moral se origina en una impresión de algo que produce agrado o desagrado en el agente. Este sentimiento moral –y no las *ideas* de la virtud o del vicio, en los términos de Hume- se experimenta repetidamente, a lo largo de la vida del agente. Al repetirse esta impresión surge, de acuerdo con las determinaciones naturales de la mente, la idea correspondiente. Esta idea es objeto de atención por parte de los agentes humanos y llegan a *creer* que representa algo distinto a una impresión de agrado o desagrado. Según hemos visto, para Hume esta creencia es lo que parece estar detrás de que se considere que la razón puede operar sobre tales ideas, supuestamente morales (cf. T 3.1.1), pero estrictamente las distinciones morales entre el vicio y la virtud se derivan solo de ciertas impresiones y, en todo caso, solo *representan* a tales impresiones (cf. Fate Norton 2003, I 14-15). De este modo, aunque el sentimiento moral pueda estar *cargado* de un cierto contenido conceptual, su rasgo distintivo es que es una cierta impresión, de una persona o de una cierta acción, considerada bajo un punto de

vista imparcial. Con base en estas ideas, Fate Norton, sostiene que Hume defiende “una teoría causal de la percepción moral, una en la cual una consideración imparcial de la intención, el motivo, o el carácter (los principios relevantes permanentes de la mente) se dice que dan lugar a placeres y dolores «peculiares», los únicos sentimientos morales” (cf. Fate Norton 2003, 163-4). Fate Norton agrega que esta percepción se debe a un cierto «sentido interno»: “...gracias a un sentimiento interno y por medio de algún sentimiento ocasionado de modo natural al reflexionar sobre esa [persona, motivo o] acción” (cf. T 3.1.1.23 SB 466). Sin embargo, no parece que Hume haga depender a tal percepción de “un sentido interno”.

J. Taylor en su estudio sobre la belleza y la virtud en Hume muestra que él toma distancia de la idea del «sentido interno», en particular por contraste con F. Hutcheson, para explicar al sentimiento moral (cf. Taylor 2008, 275 y ss).⁸⁸ Sin embargo, las ideas de Hume para explicar el *origen* del sentimiento moral son más misteriosas y complejas, como vimos en el capítulo anterior. Además, a Hume le interesa considerar a la moral ordinaria, esto es, la *práctica* común de las relaciones morales. Éste es un rasgo importante de la filosofía moral de Hume. Su explicación sobre el sentimiento moral nos enseña mucho, en un plano meramente descriptivo, del modo en que interactúan la moral ordinaria y la psicología de los agentes. En otras palabras, Hume ha dedicado considerable atención al modo en que la moral, que reúne reglas y principios de aplicación general, se relaciona con la psicología individual la cual, desde su punto de vista, está determinada por ciertas determinaciones naturales relativas al placer, el dolor, la percepción y la formación de ideas, así como a las creencias. La imagen general de la moral que resulta de la integración de estos elementos es así compleja, pero me parece tiene también un gran poder explicativo.

El pensamiento de Hume sobre las prácticas sociales encuentra en el origen de la sociedad la necesidad de sus miembros de “suplir sus defectos y llegar a ser igual a las demás criaturas, y aún de adquirir superioridad sobre ellas. [...] La sociedad se convierte en algo ventajoso mediante esta *fuerza, capacidad y seguridad* adicionales” (cf. T 3.2.2.3 SB 485). Estas ventajas, sin embargo, se ven contrariadas por la coincidencia de dos aspectos de diferente categoría cada uno.

⁸⁸ De acuerdo con Taylor, Hume se inscribe junto con Hobbes, Shaftesbury y Hutcheson en la llamada tradición sensualista moderna. De acuerdo con el sensualismo, los valores morales y estéticos son percibidos directamente por los sentidos humanos y no son objeto del intelecto o razón. Hume, por contraste con Hutcheson, no sostiene que tal percepción se deba a «un sentido interno», sino, a la simpatía, una cierta condición de las pasiones (cf. Taylor 2008, 273 y ss; v. infra p. 153).

Por una parte, son contrariadas por «las peculiaridades del temperamento natural» de los seres humanos, especialmente por el egoísmo. Pero Hume observa que el egoísmo por sí mismo sería insuficiente para minar la benevolencia natural que los seres humanos sienten unos por otros, empezando por los miembros de la propia familia (cf. T 3.2.2.5 SB 486-7). El segundo aspecto, por otra parte, se refiere a las «circunstancias externas» en que los seres humanos disfrutan de sus bienes. Hume distingue tres tipos generales de bienes: 1. La satisfacción interna de nuestra mente, 2. La buena disposición externa de nuestro cuerpo y 3. Disfrutar las posesiones adquiridas por nuestra laboriosidad y fortuna. Ahora bien, según Hume, nada ni nadie pueden impedir el disfrute del primer tipo de bien. También le parece que la afectación del segundo tipo de bien no reporta ninguna ventaja a quien perpetre el daño. Finalmente, observa que “sólo la última clase de bienes se ve expuesta a la violencia de los otros y puede además ser transferida sin sufrir merma o alteración; al mismo tiempo, nunca se tiene una cantidad tal de bienes que satisfagan a cada uno de nuestros deseos y necesidades” (cf. T 3.2.2.7 SB 487-488). Este tipo de bien se refiere a la propiedad privada y Hume señala que tanto la ambición humana por más bienes y una cierta escasez son factores externos a uno que podrían afectar su disfrute (sobre la escasez de bienes, cf. IPM 3 SB 183 y ss). Ahora bien, aunque podríamos no estar de acuerdo con las apreciaciones de Hume (actualmente, no somos ajenos a los daños del primero y segundo tipo de bienes, por ejemplo mediante el estrés, la tortura psicológica o el robo de órganos internos),⁸⁹ no veo ninguna objeción de peso para aceptar su punto general, a saber: aun cuando vivir en sociedad tiene ventajas para los agremiados en ella, alcanzar y disfrutar estas ventajas tienen que superar a la coincidencia del egoísmo natural con cierta propensión humana a despojar a otros de sus bienes, dadas condiciones de escasez y ambición, es decir, que el egoísmo natural de los seres humanos, por sí mismo no tan grave, se ve potenciado a grados dañinos en condiciones sociales de escasez de bienes y por la ambición de los mismos. En estas condiciones, que ponen en peligro las supuestas ventajas de la vida en sociedad, se requiere de normas o regulaciones para contener el egoísmo y la ambición, además de garantizar que los agremiados bajo el acuerdo social puedan disfrutar de los bienes del tercer tipo, sus propiedades. Llegamos así a la justicia y al cumplimiento de las promesas, que parecen los dispositivos normativos más importantes al respecto. Hume se

⁸⁹ Seguramente Hume tenía también conocimiento de los daños causados a “nuestra mente” por persecución política, abuso psicológico, tortura, etc.

concentra en explicar su operación y surgimiento. Al mismo tiempo, creo que algunas de las ideas más interesantes de Hume sobre la moralidad se desarrollan en torno a estas cuestiones.

Para Hume, las condiciones anteriores equivalen a decir que las determinaciones naturales que despiertan el sentimiento moral son insuficientes para superar las adversidades de la vida social, de ahí que se requieran de recursos convencionales o artificiales para remediar sus limitados poderes en condiciones de escasez relativas a la ambición humana:

...la justicia debe su origen a las convenciones humanas, y que éstas se proponen como remedio de algunos inconvenientes debidos a la coincidencia de ciertas cualidades de la mente humana con la situación de objetos externos. Las cualidades de la mente son el egoísmo y la generosidad limitada, y la situación de los objetos consiste en su facilidad de cambio, unida a su escasez en comparación con las necesidades y deseos de los hombres (T 3.2.2.16 SB 494).

En este sentido, ni la justicia ni el cumplimiento de las promesas (p.e., comerciales) son «naturales», pero en la medida en que son expresiones del sentimiento moral, el placer o aprobación que se experimenta por ellos surge “gracias a un artificio o proyecto debido a las circunstancias y necesidades de los hombres” (T 3.2.1.1 SB 477). De este modo, llegamos a la distinción entre virtudes naturales y virtudes artificiales hecha por Hume en el *Tratado*.⁹⁰

Para Hume, existen ciertas tendencias de la naturaleza humana que se manifiestan en el placer de hacer el bien a otros, o ayudarlos, o, al menos, en refrenar nuestros impulsos egoístas. Hume identifica a esta tendencia como *virtudes naturales* y cree que son anteriores al sentido del deber. Estas virtudes comprenden la afabilidad, la beneficencia, la caridad, la generosidad, la clemencia, la moderación, y la equidad según la enumeración de T 3.3.1.11 SB 578. El cuidado de los hijos por parte de los padres y la ayuda a los “miserables” muestran, según Hume, cómo operan las virtudes naturales. Desde luego, los padres conocen cuáles son sus deberes para sus hijos, “pero siente también una inclinación natural hacia ello”. Este mismo sentimiento natural motiva que ayudemos a los “miserables”. Comúnmente, sabemos que tenemos el deber de ayudar a personas en desgracia, pero saber esto no es suficiente. Lo que realmente nos mueve a ayudar es un “nuestra humanidad” (T 3.2.5.6 SB 518-519). Según Hume, ninguna acción podría sernos requerida como un deber u obligación “[a] menos que en la naturaleza humana se halle implantado algún motivo o pasión impulsora capaz de producir” tal acción (cf. T 3.2.5.6 SB 518). En otras palabras, la *idea* de deber, por sí misma, es motivacionalmente inerte y sólo logra fungir como un

⁹⁰ En la IPM Hume no recurre a esta distinción, pero su argumentación parece ser la misma que desarrollaré en los siguientes párrafos, con la cual pretende probar que las “virtudes sociales” requieren de una belleza natural y amabilidad antecedentes para que puedan lograr su cometido a favor de la utilidad pública, cf. p.e., 5.5 SB 214.

motivo para actuar cuando existe una disposición natural para respaldarlo. Además, es la *simpatía* por los otros, por su situación, el principio que nos hace considerarlos como si fuéramos nosotros mismos (cf. T 3.3.1.11 SB 579). Por esta razón, las virtudes naturales se activan y logran un efecto positivo, “bueno”, en “cada acto particular”, porque es por simpatía como son despertados en nosotros esos motivos naturales (cf. T 3.3.1.12 SB 579). Por contraste, las *virtudes artificiales*, como la justicia y el cumplimiento de las promesas, surgen por algún acto de convención, un artificio, de modo que no son propias de la naturaleza humana. En otras palabras, los actos de justicia o el cumplimiento de las promesas no son disposiciones naturales. Surgen, según desarrollaré en los siguientes párrafos, del reconocimiento tácito de sus ventajas para la sociedad y por simpatía adquieren el sentimiento moral (cf. T 3.3.1.10 577-578), pero en realidad son un acto de convención “de la humanidad en un esquema o sistema de conducta general” y esto explica que, en muchos casos, puedan incluso oponerse a los sentimientos naturales. Al igual que nosotros, Hume observa que en muchos casos “[l]os jueces quitan el dinero del pobre para dárselo al rico, dan al disoluto lo que el laborioso ha producido, y ponen en mano del vicioso los medio de causar daño, tanto a sí mismos como a otros”. La simpatía, en estos casos, nos produce una inclinación favorable hacia estas personas, casos particulares, pero este sentimiento natural es revertido porque también aceptamos a la justicia porque, “considerado en su conjunto, el esquema de la ley y la justicia es beneficioso para la sociedad” (cf. T 3.3.1.12 SB 579). De este modo, es patente que las virtudes artificiales constituyen reglas que hacen factible a la sociedad y que complementan a las disposiciones naturales al respecto. La idea general de Hume detrás de las virtudes naturales y las artificiales es “*que ninguna acción puede ser virtuosa, o moralmente buena, a menos que exista en la naturaleza humana algún motivo que la produzca, que sea distinto al sentimiento de la moralidad de la acción*” (T 3.2.1.7 SB 479, énfasis de Hume; Fate Norton 2000 I83). Esto es, que aun cuando las virtudes artificiales como la justicia se expresen o sean recibidas por los agentes como deberes morales, no serían capaces de motivar ninguna acción si no existieran disposiciones naturales que, por sí mismas, originan los motivos morales. Los intereses de Hume sobre estos temas son realmente ambiciosos.

En primer lugar, se propone explicar cómo las determinaciones naturales detrás del sentimiento moral son insuficientes para regular la vida social, de modo que pretende explicar qué «artificios» o convenciones refuerzan a las virtudes artificiales, como la justicia. Posteriormente, explica cómo para mostrar que la conducta se adecue a tales virtudes se requiere del desarrollo de

motivos moralmente significativos y, finalmente, cómo nuestros sentimientos de aprobación y desaprobación son despertados consistentemente por las acciones realizadas bajo la dirección de las virtudes artificiales (cf. Fate Norton 2000, 182; Hume no presenta en este orden estos temas. Yo seguiré este orden, derivado de la interpretación de Fate Norton). En el desarrollo de estos temas voy a seguir tanto sus ideas del *Tratado* como de la *Investigación sobre los principios de la moral*, porque a pesar de la consistencia en estrategias argumentativas hay también ciertas diferencias de interés, especialmente sobre el último de los temas mencionados.

5.2. Cómo los «artificios» o convenciones refuerzan a las determinaciones morales naturales

Respecto de la primera cuestión, cómo surgen las convenciones que dan origen a las virtudes artificiales, Hume nos dice que las ventajas de vivir en sociedad no pudieron ser alcanzadas por «la fuerza del estudio y la reflexión», esto es, que el acuerdo para establecer las condiciones que permitieran disfrutar de tales ventajas no pudo ser alcanzado racionalmente, ni por consideraciones teóricas ni por alguna reflexión instrumental a favor de ellas. Según él, estas ventajas se hicieron manifiestas a la conciencia de los agentes al buscar satisfacer el natural apetito sexual, que reúne a los seres humanos y, en vista de «la prole común» los reúne y crea vínculos entre ellos. Afirma que en esta situación, “...la costumbre y el hábito actúan sobre la delicada mente de los hijos, haciendo que se den cuenta de las ventajas que pueden obtener de la vida en sociedad y conformándoles gradualmente para ella, limando asperezas y corrigiendo las afecciones adversas que pudieran impedir su unión en sociedad” (T 3.2.2.4 SB 486). Como en el caso de la experiencia y la formación de creencias sobre cuestiones de hecho, en esta situación «el hábito y la costumbre» realizan una función regulativa y normativa de la conducta. Esta función surge por la comodidad y el placer derivados de las pasiones, lo que determina a los hombres a amoldar paulatinamente su conducta a las condiciones que favorecen el surgimiento de las ventajas de vivir en sociedad. En el caso de la justicia, cuya función normativa es el respeto a la propiedad de otras personas, su origen parece apuntar a que una sociedad funcional requiere del reconocimiento de la ventaja de respetar a la propiedad privada, el tercer tipo de bienes mencionados arriba y el más vulnerable a los abusos (cf. T 3.2.2.8 SB 488). Pero esta reflexión presupone ya a la justicia funcionando en una sociedad “como es ahora conocida” y contribuyendo al sostenimiento de las relaciones interpersonales, mediante el respeto a la propiedad privada. ¿Podemos entender el acto de convención tácita que dio origen a ella? La estrategia argumentativa

de Hume enfrenta aquí una cierta circularidad, porque la justicia requiere una cierta comprensión de que se ha hecho un daño y, sobre todo, que este daño «es una inmoralidad o un vicio cometido contra otra persona», y la percepción de esta condición parece requerir que el pensamiento moral haya sido *ya* moldeado por la influencia de la sociedad –y este requisito no puede satisfacerse en la explicación sobre el *origen* de la justicia precisamente porque no hay todavía una sociedad; más brevemente: la justicia explica ciertos rasgos fundamentales del funcionamiento de la sociedad, pero no podemos entender a la justicia sin la sociedad, de modo que una explicación que busca determinar su origen enfrenta esta dificultad.

Además, como la justicia es una virtud moral, una de sus características es que mediante ella los agentes adoptan un punto de vista imparcial y general, tanto sobre las condiciones de su aplicación como sobre los agentes involucrados. De modo que la justicia funge también como una *regla general* que engloba dicha capacidad. Este punto de vista se forma socialmente y, en consecuencia, la vida social moldea al pensamiento moral en este sentido:

...es manifiesto que en la estructura original de nuestra mente la atención más intensa está centrada en nosotros mismos; la siguiente en intensidad se dirige a nuestras relaciones y conocidos; tan sólo la más débil alcanza a los extraños y a las personas que nos son indiferentes. Por tanto, esta parcialidad y esta desigual afección [de la justicia] deberán tener influencia no sólo sobre nuestra conducta y comportamiento en sociedad, sino incluso sobre nuestras ideas de virtud y vicio, de modo que nos lleven a considerar cualquier transgresión notable de este grado de parcialidad... (T 3.2.2.8 SB 488)

La dificultad, entonces, para explicar el origen de la justicia es que su *operación* como dispositivo moral requiere de acuerdos explícitos y complejos desarrollados a partir de un cierto punto de vista que sólo parece surgir dentro de la sociedad. De modo que si la justicia no es, por una parte, una virtud natural (como hemos visto antes), esto es, una disposición natural y requiere de acuerdos complejos que sólo pueden alcanzarse dentro de la propia sociedad, surge entonces la dificultad de cómo explicar el acto de convención que le da origen.

La solución de Hume, en primer lugar, reconoce que: «Todos los miembros de la sociedad buscan establecer estabilidad en la posesión de los bienes externos, dejando que cada uno disfrute pacíficamente de aquello que pudo conseguir gracias a su laboriosidad o suerte» (cf. T 3.2.2.9 SB 489). Posiblemente, este antecedente motivó a los miembros de la sociedad primitiva a buscar un acuerdo a favor de tal interés. Sin embargo, este acuerdo no puede ser como una promesa y no tiene, en estas condiciones, ni siquiera que ser explícito para los agremiados originales. En efecto, según Hume: “Esta convención no tiene la naturaleza de una *promesa*, pues hasta las mismas

promesas... surgen de las convenciones humanas” (cf. T 3.2.2.10 SB 490). Buscando la solución de estas dificultades, Hume evita incurrir en petición de principio, porque la justicia no podría surgir por un acto de convención explícito. En las condiciones en que suponemos que surgió ella misma tendría que ser parte de las condiciones que harían posible tal acto de convención. En otras palabras, en las condiciones sociales primitivas no existían los recursos normativos para hacer promesas ni, en consecuencia, para establecer acuerdos que restringieran la propia conducta mediante alguna fórmula o dispositivo auto-regulativo, por lo tanto, la justicia no puede surgir por un acuerdo semejante al de una promesa. La solución de Hume a esta dificultad es la siguiente:

La convención consiste únicamente en un sentimiento general de interés común: todos los miembros de la sociedad se comunican mutuamente este sentimiento, que les induce a regular su conducta mediante ciertas reglas. Yo me doy cuenta de que redundará en mi provecho el que deje gozar a otra persona de la posesión de sus bienes, dado que esa persona actuará de la misma manera contigo. También el otro advierte que una regulación similar de su conducta le reportará un interés similar. Una vez que este común sentimiento de interés ha sido mutuamente expresado y nos resulta conocido a ambos, produce la resolución y conducta correspondiente (cf. T 3.2.2.10 SB 490).

Según Brand, de este acto de convención surge la regla general de justicia, la cual da “estándares normativos para la interpretación de la experiencia inmediata y para las evaluaciones morales naturales, «incultas»” (cf. Brand 1992, 38). El estándar normativo, referido a evaluaciones incultas, es un tipo de acuerdo o convención no explícita ni que necesite formularse verbalmente. Así, Hume nos ofrece una explicación plausible de cómo surgen convenciones en condiciones pre-rationales y, aún, previas a los dispositivos morales que sancionan tales acuerdos. El reconocimiento del interés común, por parte de cada agente individual, es todo lo que se necesita, según ilustra estupendamente: “Cuando dos hombres impulsan un bote a fuerza de remos lo hacen en virtud de un acuerdo o convención, a pesar de que nunca se hayan prometido nada mutuamente” (T 3.2.2.10 SB 490). Hardin ha llamado la atención sobre este pasaje y afirma que es un antecedente de las ideas David Lewis 1969 sobre la convención (Lewis, 1969). Según Hardin, Hume soluciona aquí el problema de cómo explicar la coordinación entre agentes que o no tienen aún lenguaje o no pueden recurrir a él por alguna razón (v. Hardin, 82; Rescorla, §3.1 “Hume”). La teoría general de la moralidad de Hume, me parece a mí, encuentra un importante respaldo en este argumento a favor del acuerdo o convención que explican a la justicia.⁹¹ En la cita anterior Hume reconoce que eventualmente este sentimiento de interés compartido y convenido implícitamente como una regla general se expresa y resulta conocido para los

⁹¹ Posiblemente, Hume sea el primer autor en usar al concepto de convención para explicar el origen de la moralidad, v. p.e. T 3.2.1.11 SB 481.

agremiados sociales. Según su explicación genealógica, esto se logra a través de un proceso gradual, el cual logra asentar y fortalecer la regla de justicia en la sociedad y en la conciencia de los agentes: "...la regla de la estabilidad de la posesión, por surgir gradualmente e ir adquiriendo fuerza mediante una lenta progresión, y porque experimentamos repetidamente los inconvenientes que resultan de transgredirla" (cf. T 3.2.2.10 SB 490). Se puede explicar así cómo surgen las convenciones que dan origen a las reglas morales.

Con todo, parece haber un cierto titubeo en el modo en que según Hume las reglas morales son establecidas. Nos dice, por ejemplo, que las reglas de propiedad son establecidas por su utilidad pública, aunque también pudieron haber sido «fijadas por la imaginación» (T 3.2.3.4 SB 504, v. nota 71). La razón para esta duda parece ser, por una parte, que la idea misma de propiedad tiene que surgir mediante algún *artificio* en la mente, porque no hay en la naturaleza, en «el imaginario estado de naturaleza que precede a la sociedad», ninguna marca que indique que tal o cual bien pertenece a determinada persona (tampoco hay ninguna necesidad de la justicia). Por lo tanto, no pueden ser la razón ni las impresiones las que den origen a la idea de propiedad, sino la imaginación (cf. T 3.2.2.28 SB 501). La idea de propiedad también puede surgir de la noción de utilidad social y Hume parece pensar que es esta noción la que pudo haber despertado a la imaginación. Si analizamos la idea de propiedad esto parece plausible, porque ésta es una idea que fue ganando en complejidad –conforme su uso se fortalece socialmente– y que llega a comprender a las reglas más particulares de *ocupación, prescripción, accesión y sucesión* (cf. T 3.2.3.5 SB 505). Voy a considerar brevemente la relación que guarda, según Hume, la comprensión de estas reglas con la noción de utilidad.

La primera de estas reglas, según Hume, es la más importante para el establecimiento de la sociedad y garantiza la estabilidad e invariabilidad en las posesiones de sus miembros. Por "posesión" Hume entiende aquella relación entre un agente y un bien (en sentido amplio) tal que el agente puede «tocar directamente a aquel bien y está en posibilidad de utilizarlo, moverlo, alterarlo o destruirlo de acuerdo con el placer o provecho del momento» (cf. T 3.2.3.7 SB 506). Su importancia reside en que esta regla "...es la razón por la que, mediante un acuerdo universal, los hombres se restringen unos a otros por lo que ahora llamamos reglas de justicia y equidad" (cf. T 3.2.3.6 SB 505). De acuerdo con esto, la regla de la justicia se deriva de la regla de propiedad.

En efecto, el uso primordial de la justicia puede entenderse como la necesidad de hacer valer la relación de propiedad. En la teoría genealógica de la moral de Hume el crecimiento de las sociedades requirió que las normas de propiedad y de justicia fueran cada vez más complejas. Este proceso no es racional sino que fue impulsado por la costumbre y el hábito, que fortalecieron en los agentes la percepción de utilidad y de placer asociados con tales reglas (v. p.e., T 3.2.2.4 SB 486). Eventualmente, la relación de posesión primera, instaurada como derecho, se va desvaneciendo y se requiere de algo más fuerte para mantener el derecho inicial de posesión. Hume llama *prescripción* al dispositivo regulativo “que confiere a la persona suficiente título de propiedad sobre cualquier cosa que posea”. La función de esta regulación es combatir el efecto del tiempo y el olvido. Hume no lo menciona explícitamente, pero la *prescripción* parece requerir de algunos medios para garantizar el derecho a la propiedad y con ello resolver las «controversias» que pudieran surgir (cf. T 3.2.3.9 SB 507-508). Hoy hablaríamos de esta regulación en términos de “título de propiedad”, entre otros nombres. Las restantes dos reglas parecen derivarse de ésta. Una vez que tenemos “clara y distintamente” acreditada la propiedad de un bien, tenemos derecho a aquellos objetos «íntimamente conectados» con dicho bien, como «los frutos de nuestro jardín, las crías de nuestro ganado», etc.⁹² Hume llama a esta regla “accesión” [*accession*] (cf. T 3.2.3.10 SB 509).⁹³ Por último, la regla o derecho de sucesión consiste en que la posesión de alguna persona pase a «quien o quienes le son más queridos». Esta regla reúne tanto elementos naturales como a la propia utilidad o «interés general de la humanidad». Respecto de los primeros, Hume señala que a la mente humana le es natural admitir la relación de sucesión a los seres «más queridos» y esta relación, de por sí natural en el caso de los hijos, puede verse «ayudada por la influencia de la asociación entre ideas, la cual inclina naturalmente nuestra atención hacia los descendientes». Especialmente, cuando el titular de un bien ha fallecido, la idea de los descendientes, nuevamente por asociación de un objeto inmediato con otro, como propietarios se «consolida» en la mente (cf. T 3.2.3.11 SB 510-513). En este caso, las determinaciones naturales más básicas de la mente complementan y fortalecen este proceso de sucesión.

La operación de sucesión es, en términos de Hume, simbólica, en el sentido de que es una fantasía, puesto que no hay ninguna marca o propiedad de los objetos que sea transmitida del

⁹² Lamentablemente, el siguiente ítem de esta lista es: “el trabajo de nuestros esclavos”.

⁹³ En el derecho, accesión se refiere al dodo de adquirir el dominio, según el cual el propietario de una cosa hace suyo, no solamente lo que ella produce, sino también lo que se le une o incorpora por obra de la naturaleza o por mano del hombre, o por ambos medios a la vez, siguiendo lo accesorio a lo principal. Hume sigue este uso.

dueño anterior al nuevo. Toda la operación sucede en la imaginación, la cual es ayudada por la vivacidad de alguna idea y alguna relación con otra impresión (la misma estructura que explica a las creencias causales, v. supra §1.2). Así, la idea de propiedad original es muy vivaz (Hume incluso la califica, como vimos antes, como *clara y distinta*) y la imaginación la relaciona con el nuevo titular de la propiedad, aquel a quien se le ha sucedido la propiedad. De esta persona tenemos una impresión presente, de modo que la imaginación fácilmente relaciona a la vivaz idea de propiedad y la *transfiere* a esta persona. Así, la mente (Hume no nos dice de quién es esta mente, pero suponemos que toda la operación descrita funciona igual para todo aquel que comparta estos elementos) es engañada y cree que ha tenido lugar «la misteriosa transmisión de propiedad» (cf. T 3.2.5.2 SB 515) Ahora bien, aunque estamos en el terreno de la imaginación, esta operación simbólica, de la imaginación, tiene consecuencias normativas y motivacionales relativas a la utilidad. Hemos visto antes cómo las reglas generales de la justicia se derivan de la idea de propiedad y aparece como una institución moral necesaria para el funcionamiento de la sociedad (v. T 3.2.2). Este rasgo es parte de su utilidad. Consideremos ahora a la promesa.

El complejo derecho de propiedad presentado arriba hace surgir ciertas obligaciones entre los miembros de la sociedad. Éstas se pueden presentar de varias maneras y me parece que cada una de estas formas articula instituciones sociales. Por ejemplo, el derecho de titularidad de propiedad, o de *prescripción* en términos de Hume, da lugar a las instituciones de contrato, de legitimidad de tal derecho (magistrados, árbitros, etc.) y otras instituciones para garantizarlo. Estas instituciones asumen la forma de contratos o promesas, esto es, los receptores de la obligación *prometen* respetar la titularidad de la propiedad. Podemos hacer consideraciones similares para la sucesión de propiedad. Ahora bien, mientras que el titular de la propiedad tiene derecho sobre la misma, otros agentes, individuales o institucionales, tienen obligaciones relativas al respeto de dicho derecho. Desde la perspectiva de Hume, como estas obligaciones se derivan o son correlativas a la ficticia o simbólica relación de propiedad, la propia obligación es igualmente ficticia o simbólica, esto es, tiene lugar sólo en la imaginación de los agentes.

Hume cree que la obligación detrás de las promesas es artificial, esto es, no son reglas morales que surjan de determinaciones naturales, sino que surgen por, y son inteligibles como, convenciones humanas. Además, cree que aun cuando por un acto de convención se haya hecho alguna promesa, su carácter obligatorio no surge de este acto (cf. T 3.2.5.1 SB 516). La primera afirmación se deriva del razonamiento de que si las promesas fueran algo natural e inteligible, al

pronunciar la fórmula *yo prometo* (o alguna otra similar) algún acto mental debería acompañarla y la obligatoriedad se tendría que derivar de tal acto. Los actos que Hume considera al respecto son la resolución, el deseo y la volición. Pero ninguno de ellos parece producir el efecto normativo buscado. El primero, por sí mismo, no nos impone ninguna obligación. En efecto, tener la resolución de llegar temprano o de dejar de fumar no producen ninguna obligación en el agente, excepto, quizá, una cierta actitud al respecto. El deseo tampoco produce obligación alguna porque muchas veces asumimos obligaciones sin tener el deseo correspondiente o incluso sin desear asumirla. Como cuando alguien promete reunirse con alguien más a una hora que detesta por alguna razón. Finalmente, la volición no puede producir ninguna promesa porque tiene influencia sólo sobre acciones presentes (Hume considera que la volición es la fuerza o deseo final que acompaña a una acción, ref.) y las promesas se refieren siempre a un tiempo futuro. De estas consideraciones Hume concluye que la obligación de cumplir una promesa “tendrá que consistir, necesariamente, en la *volición de la obligatoriedad misma* que se manifiesta en la promesa”. Es decir, que la obligación surge sólo de nuestra mera voluntad y deseo, pues no hay ningún acto mental que nos constriña. Ahora, Hume se pregunta si concebir a la obligación de la promesa de este modo no es «un absurdo manifiesto» (cf. T 3.2.5.2-3 SB 516-517). La respuesta de Hume a esta pregunta, me parece a mí, es del todo consistente con la explicación genealógica sobre la moralidad y la motivación moral.

En la medida en que está considerando a la obligación moral de una promesa, Hume recuerda en primer lugar que la moralidad depende por completo de nuestros sentimientos, particularmente de un cierto sentimiento de agrado cuando aprobamos moralmente algo y de un cierto sentimiento de desagrado cuando rechazamos moralmente algo. De modo que el surgimiento de una obligación tendrá que equivaler al surgimiento de un cierto sentimiento, pero la dificultad al respecto es que a cada uno de los agentes humanos «cambiar naturalmente sus sentimientos les resulta tan imposible como cambiar los movimientos celestes», de modo que el acto de la voluntad de prometer no puede producir el sentimiento de obligación correspondiente (T 3.2.5.4 SB 517). Parece seguirse de este argumento que el sentimiento de obligación moral no puede surgir por un acto *interno* a la mente del agente, sino que tiene que provenir de factores compartidos socialmente, cuando el agente es miembro de un *colectivo*, como comenta Wiggins (v. Wiggins 2006, 49). De estas ideas se sigue que para Hume la obligación de cumplir las promesas tendría que ser una obligación social, cuya fuerza constrictiva sobre el agente proviene

de factores sociales o *externos* al mismo, tal como acepta en: “...me atrevo a concluir que las promesas son invenciones humanas basadas en las necesidades e intereses de la sociedad” (T 3.2.5.7 SB 519). Con estos antecedentes, examinaré el segundo tema que me ocupa en este capítulo: cómo los agentes conforman su conducta con tales reglas.

5.3. Cómo los agentes conforman su conducta a las reglas morales

Siguiendo a Fate Norton, para responder a dicha cuestión se requiere que los agentes desarrollen motivos moralmente significativos, lo que, dentro de la teoría de Hume, equivale a sostener que se requiere que el sentimiento de utilidad social se experimente con placer o desagrado, de acuerdo con su cumplimiento o incumplimiento. Hume logra con esta explicación, si es correcta, mostrar cómo se relacionan elementos de la psicología individual con las normas sociales relativas a la justicia y la utilidad. En esta cuestión, una vez más, las determinaciones naturales que operan sobre los materiales de la mente son muy relevantes para la explicación de Hume. Aun así, se puede presentar de varias maneras la dificultad en torno a esta cuestión. Por ejemplo, de acuerdo con las ideas de Hume, quizá nada sea más contrario al interés público que la injusticia y la deshonestidad, pero, al mismo tiempo, “el interés público no está ligado por naturaleza a la observancia de las reglas de justicia [...] los hombres no se preocupan en su conducta ordinaria por algo tan lejano como el interés público cuando pagan a sus acreedores, cumplen sus promesas o se abstienen de robar, saquear o de cometer injusticias de cualquier tipo” (T 3.2.1.11 SB 480-481). En otras palabras, aun cuando los agentes puedan actuar de acuerdo con el “deber y obligación” de los que hablan los moralistas, seguramente sus motivos no son necesariamente “el amor a la humanidad”. En efecto, en las transacciones de la vida ordinaria rige el interés propio, la buena fama (volveré un poco más abajo a esta idea), de modo que el motivo originario de la justicia no es el “amor universal, o respeto por los intereses de todos los hombres” (cf. T 3.2.1.12-13 SB 481-482). Esta observación de Hume parece conducir a una cierta circularidad en su teoría general sobre la moralidad.

En los pasajes citados, Hume afirma que el *motivo* por el que los agentes actúan es sobre todo el interés propio, aun cuando tales motivos coincidan con el deber y la obligación característicos de las acciones justas –como pagar las deudas contraídas, así sea a usureros.⁹⁴ Pero

⁹⁴ Hume, en efecto, después de observar que censuramos a un padre por no atender a su hijo porque «carece de una afección natural, el deber natural de todo padre», concluye que este deber está determinado por un motivo natural, el cual es distinto al deber mismo. Su idea de que los motivos morales están necesariamente detrás de las acciones

hemos visto que el sentimiento moral presupone un punto de vista imparcial sobre otros agentes. Así, el sentimiento moral, cuando funge como motivo de la acción, no es auto-interesado, sino benevolente o generoso, esto es, está orientado hacia al interés de algún otro agente. Hume enfatiza, en este contexto, que la naturaleza del motivo es lo que distingue a las acciones como morales: "...ninguna acción puede ser virtuosa sino en cuanto procede de un motivo virtuoso. Luego el motivo virtuoso deberá preceder al respeto por la virtud: es imposible que el motivo virtuoso y el respeto por la virtud puedan ser la misma cosa" (T 3.2.1.9 SB 480). Me parece que hay dos formas de alejar la circularidad entre estas ideas.

En primer lugar, como hemos visto, Hume concibe a los agentes motivados tanto por el egoísmo como por la benevolencia. Sin embargo, los motivos benevolentes tienden a contrarrestar a los egoístas, de modo que las prácticas morales y sociales compartidas son adecuadas para que los motivos benevolentes den lugar a acciones virtuosas. Éste es uno de los lugares comunes en la IPM (cf. Fate Norton 2003, I61). Muy plásticamente, Hume presenta del siguiente modo esta pluralidad de motivos: "Junto con los elementos del lobo y de la serpiente yace en nuestra naturaleza alguna chispa de amistad por la humanidad, alguna partícula de la paloma..." (Cfr. IPM SB 271). En consecuencia, podemos responder a la objeción de inconsistencia mencionando los distintos motivos que están detrás de las acciones humanas: éstos no siempre serán benevolentes y es posible que los motivos egoístas se sirvan de la apariencia de la benevolencia. En segundo lugar, precisamente esta condición y la pluralidad de motivos detrás de las acciones humanas es lo que suscita la cuestión que ahora abordamos, esto es, qué elementos contribuyen a la explicación de cómo se desarrollan motivos moralmente significativos. Simplificando de esta manera, resalta entonces que la inconsistencia en cuestión está en el núcleo de los retos a los que la teoría moral de Hume se dirige.

La respuesta de Hume parece ser que las determinaciones naturales son reforzadas por convenciones sociales, las cuales expresan las reglas que nos interesa explicar, como la justicia. Posteriormente, estas convenciones son a su vez percibidas como cierto sentimiento placentero o displacentero por parte de los agentes. Sobre este último aspecto, hay ciertas diferencias en la explicación de Hume. Mientras que en el *Tratado* tal sentimiento depende de la simpatía, en la

morales la recoge en la siguiente «máxima indudable»: "...ninguna acción puede ser virtuosa, o moralmente buena, a menos que exista en la naturaleza humana algún motivo que la produzca, que sea distinto al sentimiento de la moralidad de la acción" (T 3.2.1.7 SB 479, énfasis de Hume).

IPM parece haber ciertos sentimientos naturales básicos que, aunados a la belleza y amabilidad de las acciones virtuosas, refuerzan a las convenciones morales. Pero la estrategia argumentativa es consistente en ambas obras. Hume en el *Tratado* describe este proceso de ampliación de los sentimientos morales a partir de las virtudes naturales hasta cubrir a las virtudes artificiales (esto es, desde las condiciones de la moral natural hasta las complejas manifestaciones de la moral), como un *progreso* de los sentimientos, el cual es ayudado por factores sociales, como «el artificio de los políticos»:

Aunque este progreso de los sentimientos sea natural e incluso se produzca de un modo necesario, es cierto que se ve ayudado por el artificio de los políticos, que, para gobernar a los hombres con mayor facilidad y conservar la paz en la sociedad, se han esforzado por inculcar aprecio por la justicia y aborrecimiento por la injusticia (T 3.2.2.25 SB 500).⁹⁵

La idea básica de este proceso progresivo es que los agentes humanos están dotados de ciertas determinaciones naturales que les suscitan motivos morales bajo la forma de un cierto sentimiento placentero o doloroso. Hemos visto que estas determinaciones básicas no comprenden motivos para actuar de acuerdo con la justicia ni con la idea de utilidad social, y que el origen de éstos presupone algún acto «artificial», como la convención (cf. T 3.2.2.21 SB 496). La dificultad que nos interesa, en este punto, puede expresarse así: ¿cómo logra la consideración de justicia o de utilidad social motivar a agentes que por naturaleza no sienten ninguna inclinación al respecto? Hume, además, considera que la justicia es un sentimiento que “no está basado en la razón”, porque no presupone “el descubrimiento de ciertas conexiones y relaciones de ideas eternas, inmutables y universalmente obligatorias” (T 3.2.2.20 SB 496), de modo que, en sus términos, el problema de cómo tal sentimiento llega a motivarnos se expresa del siguiente modo: “*Por qué unimos la idea de virtud a la justicia y la de vicio a la injusticia*”, esto es, Hume necesita explicar la obligación *moral* hacia la justicia, o sentimiento de lo justo y lo injusto (cf. T 3.2.2.23 SB 498). La respuesta de Hume reúne a las condiciones sociales con la simpatía.

Hume explicó, según vimos antes, el proceso por el cual podrían surgir reglas o convenciones morales en condiciones que excluyeran la racionalidad y la posibilidad de hacer acuerdos tácitos (v. supra §5.2). En el contexto de la cuestión que ahora nos interesa complementa esta explicación señalando que en las pequeñas sociedades primitivas los agentes reconocen las ventajas de seguir las reglas morales y restringir a las pasiones del egoísmo y la generosidad limitada. Así, dice que los agentes, “en un principio se ven inducidos a imponerse y obedecer estas

⁹⁵ El libro de Baier *A Progress of Sentiments...* (1991) toma su nombre de este pasaje.

reglas, tanto en general como en cada caso particular, únicamente por respeto a su interés”. Los agentes aprecian directamente las ventajas de actuar así y las ventajas que les reporta compensan someter el egoísmo. Sin embargo, conforme crecen las sociedades hasta “convertirse en tribus o naciones” ese interés particular “se hace remoto [y] los hombres no perciben ya con facilidad que cada vez que se quebrantan estas reglas se sigue el desorden y la confusión...”. En sociedades grandes, de acuerdo con Hume, los agentes pierden más fácilmente de vista las ventajas de seguir las reglas de justicia y utilidad y no ven que quebrantarlas va en contra de su propio interés. Pero en este punto la *simpatía*, una determinación natural, contribuye a fortalecer a las reglas artificiales. Según la explicación de Hume, cuando presenciamos en estas condiciones los padecimientos de quienes sufren injusticias o condiciones adversas viviendo en sociedad, entonces por «simpatía participamos del malestar del afectado» y por extensión del malestar que experimentamos ante ello, «llamamos vicio a lo que produce tal malestar y virtud a lo que produce satisfacción», y agrega que “ésta es la razón de que el sentimiento del bien y del mal morales siga siempre a la justicia y a la injusticia.” Pero esto describe, de otra manera, la regla general de la justicia, porque al generalizar sobre el vicio, la virtud, o correspondientemente sobre la justicia y la injusticia, adoptamos un cierto punto de vista imparcial (y recíproco) hacia los demás miembros de la sociedad: “*La regla general* va más allá de los casos particulares de que surgió, mientras que al mismo tiempo *simpatizamos* con los demás en los sentimientos que de nosotros tienen (cf. T 3.2.2.24 SB 499). Así, como hemos visto antes siguiendo a Brand, la regla general permite a los agentes hacer interpretaciones normativas de su experiencia inmediata. Éste es un rasgo fundamental de la experiencia moral, tal como parece seguirse de la descripción que hace Hume de cómo interiorizamos las reglas morales. No hay que olvidar que para él la simpatía es una determinación natural que fortalece a las convenciones sociales.

Como hemos visto en el tratamiento de las pasiones, la simpatía no es ella misma una pasión, sino un estado mental que funciona como mecanismo para la transmisión de las pasiones (cf. T 2.1.11.2-8 SB 316-320; Merrill, 266). Su papel es hacer posible la percepción de ciertas situaciones y eso provoca en los agentes respuestas emocionales de tristeza, enojo, alegría, o, como en el caso de la percepción de injusticias que ahora consideramos, sentimientos de dolor o displacer que catalizan, posiblemente, emociones morales como indignación, enojo, retribución, etc. Ahora bien, con base en estas ideas y en el papel de la simpatía a favor de las convenciones sociales, Hume desarrolla una estrategia argumentativa a la que recurre también en la *IPM*. Parece

seguirse de las consideraciones anteriores que *sin* la simpatía ni la justicia ni el interés público lograrían motivar a los agentes a actuar de acuerdo con ellas. Además, de acuerdo con ideas de Hume, ningún recurso «artificial» o ninguna institución social, como la política, podría despertar los motivos adecuados, según se desprende del siguiente pasaje:

Cualquier artificio de los políticos puede ayudar a la naturaleza en la producción de los sentimientos que ésta nos sugiere, y en alguna ocasión puede incluso producir, por sí solo, aprobación o aprecio por una acción determinada. Pero es imposible que sea la sola causa de la distinción que realizamos entre vicio y virtud; si la naturaleza no nos ayudara en este respecto sería inútil que los políticos nos hablaran de lo honroso o deshonroso, de lo digno o de lo censurable. Estas palabras nos serían totalmente ininteligibles y no uniríamos a una idea a ellas más de lo que haríamos si pertenecieran a una lengua que nos fuera completamente desconocida (T 3.2.2.25 SB 500).

Hume, consistentemente, establece que las distinciones morales surgen necesariamente de alguna determinación de nuestra naturaleza, así se trate de las relativas a las virtudes artificiales. Otro paso de la estrategia mencionada, según esta cita, es que los términos relativos a las valoraciones sociales, como ‘honor’, ‘digno’ o ‘censurable’ –y sus opuestos– no pueden producir en nuestra mente ninguna idea si nuestra mente no tuviera alguna disposición natural para ellos. Ahora, el siguiente paso de la estrategia es sacar la conclusión de que las ideas correspondientes, derivadas directamente de nuestra naturaleza, son los términos morales valorativos básicos. En otras palabras, Hume parece decir que las valoraciones morales sólo son significativas y *funcionales* en relación con la motivación en la medida en que logran relacionarse adecuadamente con nuestra naturaleza moral básica. Las determinaciones naturales detrás de la moralidad y la motivación parecen rodear a las convenciones sociales al respecto. Los agentes humanos responden adecuadamente a las reglas sociales de la moralidad y desarrollan los motivos moralmente significativos porque por naturaleza tienen una cierta predisposición a hacerlo. Desarrollaré ampliamente esta idea, la cual Hume presenta aquí, me parece que por vez primera, pero también la encontramos en T 3.3.1.11 SB 579 y en IPM 5 SB 214-215.

Para empezar, consideraré con más detalle el proceso de interiorización de las reglas y cómo, a pesar de que es un proceso social, pronto encuentra vinculación con las determinaciones naturales. Así, según Hume:

...como los padres observan en seguida que un hombre es tanto más útil para sí mismo y para los demás cuanto mayor sea su grado de probidad y honor, y que estos principios son más poderosos cuando la costumbre y la educación van en ayuda del interés y la reflexión, por estas razones se ven inducidos a inculcar a sus hijos, desde su más tierna infancia, buenos principios, enseñándoles a obedecer las reglas por las que la subsiste la sociedad y a considerarlas como valiosas y dignas de respeto, así a como a tener su violación por algo bajo e infame (T 3.2.2.26 SB 500-501).

Hume ofrece aquí, además de una descripción de cómo interiorizamos las reglas morales, una caracterización interesante de la moralidad, en general, como un sistema regulativo que depende de las relaciones sociales y de la interacción de los diversos agentes. Ésta no es, por contraste con la moral kantiana, una moral que presupongo que un agente impersonal, aislado, autónomo, resuelva solo la cuestión de cuál es su deber o qué debería hacer en tales o cuales circunstancias. El agente humeano, según se desprende de las citas anteriores, en relación con estas preguntas, estaría interesado sobre todo en su honor, su dignidad, la vergüenza, la buena fama, etc. Según Hume, estos factores fortalecen a la justicia y al rechazo a la injusticia: “Lo que más contribuye a incrementar, después, la consistencia de esos sentimientos es el interés que tenemos por nuestra reputación, una vez que se ha establecido ya entre los hombres la opinión de *que a la justicia y a la injusticia les acompaña el mérito o demérito*” (T 3.2.2.27 SB 501). Para este agente el punto de vista de los otros, el cual comparte, marca los estándares normativos que él mismo comparte y promueve (más adelante destaco que este rasgo es crucial para entender la psicología del «observador ideal» del agente humeano en el ámbito social, v. §6.3). Al hacerlo, aun cuando no sea evidente para él, sigue al mismo tiempo las determinaciones de su naturaleza que lo predisponen a la moralidad. Con estas ideas en mente examinaré ahora el desarrollo que hace Hume de esta estrategia en la IPM. Sostendré que aun cuando es básicamente la misma estrategia, Hume aporta otros elementos de interés.

La principal divergencia en la estrategia de Hume en las IPM es que no recurre a la simpatía como la determinación que respalda a las convenciones sociales. Con todo, parte de que la utilidad social nos es placentera, pero ¿por qué es placentera o por qué nos agrada? La sección V de las IPM la dedica a responder esta cuestión. Observa, en primer lugar, que la utilidad en distintos órdenes nos resulta placentera. Según él, la utilidad es un criterio valorativo positivo de «la vida cotidiana». Así, se valora positivamente a una persona que ha demostrado su utilidad para el público, por sus servicios a la sociedad o a la humanidad. También los objetos útiles resultan placenteros, aun cuando no sean regulares ni elegantes, ya sea que se trate de un barco o un edificio o cualquier dispositivo, por desproporcionado o deforme que pueda parecernos a primera vista. Correspondientemente, una persona cuyos hábitos y comportamientos son «perjudiciales para la sociedad y con aquellos con quienes se relaciona es objeto de desaprobación y suscita fuertes sentimientos de disgusto y aborrecimiento» (cf. IPM, SB 212). A pesar de esto, nos dice Hume, “la dificultad de dar cuenta de estos efectos de la utilidad o de su contrario ha impedido que los

Filósofos los admitan en sus sistemas de ética”. La dificultad, parece ser, que no se puede explicar cuál es el *origen* de esta valoración placentera asociada con la utilidad. «Los filósofos modernos», Hume se refiere a ellos, explican por una especie de acto de convención el origen de la utilidad. Según este relato convencional, dichos filósofos han observado la importancia de la utilidad y de ahí han inferido que “todas las distinciones morales surgen de la educación y fueron inventadas en un primer momento, y después fomentadas, por la habilidad de los políticos, y ello con vistas a volver tratables a los hombres y mitigar su egoísmo y ferocidad naturales” (cf. IPM, SB 215). Hume no desestima la importancia de la educación y de los preceptos para influir positivamente sobre los sentimientos morales, pero no cree que sean suficientes para explicar el origen de las distinciones morales, esto es, «el afecto o la aversión morales». La razón principal de Hume al respecto es que ni la educación ni la habilidad de los políticos para influir sobre los sentimientos morales serían eficaces si no hubiera en la mente de los agentes morales una disposición natural para experimentar tales sentimientos:

Si la naturaleza no hubiera hecho una tal distinción, basada en la constitución original de la mente, las palabras honorable y vergonzoso, amable y odioso, noble y despreciable nunca hubieran tenido lugar en ninguna lengua, ni hubieran sido capaces los políticos, si hubieran inventado estos términos, de volverlos inteligibles o de hacerles transmitir una idea a la audiencia (IPM, SB 215).

Hume defiende aquí que la moral no surge ni por un acto de convención dirigida por la utilidad social y el interés de someter los instintos egoístas y «feroces» de los agentes humanos, sino que, en todo caso, las distinciones morales están ya inscritas en la naturaleza humana, pues forman parte de “la constitución original de la mente”. La moralidad, de este modo, se origina directamente en ciertas determinaciones de la naturaleza humana. Además, Hume indica que los términos morales *evaluativos* más básicos son, agrupados por pares de opuestos: honorable y vergonzoso, amable y odioso, noble y despreciable.⁹⁶ Es decir, de acuerdo con el contexto, que las respuestas morales naturales consisten en una reacción emocional correspondiente a dichos términos y que tal respuesta es normativa en la medida en que los agentes tienden a regular su conducta en consonancia con tales emociones, por ejemplo, buscando aquellas situaciones que les reportan honor y alejándose de las que les reportan vergüenza, etc. En conjunto, estas ideas apuntan a que no podría haber ningún acto de convención sobre prácticas que supuestamente

⁹⁶ Baier indica que son los términos morales más relevantes para Hume. De acuerdo con ella, Hume con esto es un antecesor del trabajo de A. Gibbard *Wise Choices, Apt feelings* (1990), donde consideró que términos reactivos como éstos y en particular el desdén y la vergüenza, están en el origen mismo de la moralidad, en lugar de la culpa y el enojo, que parecen más asociados con la explicación kantiana correspondiente (cf. Baier 1993, 449).

reportarán alguna utilidad social si tales prácticas no se relacionarán adecuadamente con nuestras ideas morales naturales y más básicas. Así, según la explicación genealógica de la moral que propone Hume, la educación y los preceptos de los políticos pueden ser eficaces porque se relacionan o hacen más potentes tales determinaciones naturales. En consecuencia, se sigue que la explicación que rechaza Hume ha invertido los términos al suponer que la utilidad surge por convención y que nos place reconocerla y fomentarla. En realidad, según el argumento de Hume, tenemos una disposición natural para experimentar sentimientos morales básicos bajo alguna de las manifestaciones mencionadas en la cita. Estos sentimientos, tal como los ha apareado Hume, son correspondientemente placenteros y dolorosos. El placer que despierta la utilidad es un sentimiento moral cuando lo experimentamos en relación con tales sentimientos básicos. (Esto explicaría porque el placer que experimentamos al constatar lo útil que ha resultado nuestro flamante sacacorchos no calificaría como una emoción moral.) En relación con la infidelidad y la vergüenza, Hume nos ofrece un ejemplo de cómo opera la función normativa de estas emociones básicas.

Hume considera que es necesario restringir la fuerte tentación que la infidelidad representa para las mujeres —no parece preocuparle mayormente si los varones no son fieles. La infidelidad le interesa por la necesidad que tienen los varones de alguna garantía de que los hijos que cuidan sean efectivamente suyos. La utilidad social de esto parece derivarse del hecho de que a través del cuidado de los hijos y de la familia se mantienen lazos sociales importantes. El argumento de Hume parte de la observación de que la naturaleza ha determinado que el “principio generativo” pase del hombre a la mujer, esto es, que la gestación y nacimiento de los hijos tenga por causa al hombre pero se realice en la mujer. De esto, dice, “se deriva la gran diferencia existente en la educación y deberes de los dos sexos” (T 3.2.12.3 SB 571). Estos “deberes” consisten, para el hombre, en “trabajar para el cuidado y mantenimiento de sus hijos” de lo que se sigue que sea “razonable y aun necesario darles alguna seguridad” sobre la paternidad de los hijos. Este argumento —que es por lo menos polémico— implica, según Hume, el deber de la fidelidad femenina, el único que puede garantizar que los hijos sean del padre que los cuida y mantiene. Con todo, Hume considera que la fidelidad de la mujer es un deber moral y no legal.⁹⁷ Su razón

⁹⁷ Por contraste, por ejemplo, con las ideas de Kant sobre el mismo tema. No desarrollaré este tema, aunque brevemente indicaré que en el tratamiento que hace Kant de la infidelidad recurre al lenguaje de castigo, retribución y culpa (cf. Kant 1797 6:336-337). Baier, al estudiar la posición de Kant sobre la infidelidad, señala que su tratamiento

al respecto parece débil, porque dice que es un deber moral debido a las dificultades para presentar pruebas de valor jurídico, pero esto le permite orientar la cuestión de la restricción al terreno de la moral y pregunta en consecuencia: “¿Qué restricción impondremos entonces a las mujeres para contrarrestar una tentación tan fuerte como la tienen por la infidelidad?” y responde:

No parece que sea posible otra restricción sino la de castigarlas por medio de la mala fama o reputación; en efecto, este castigo ejerce una poderosa influencia sobre la mente humana. A fin de obligar, pues, al sexo femenino a guardar la debida compostura unimos a su infidelidad un peculiar grado de vergüenza [*shame*],⁹⁸ superior al originado meramente por lo inmoral de la falta; hay que señalar que del mismo modo alabamos a la mujer por su castidad (T 3.2.12.4 SB 571).

Así, para Hume la vergüenza es un dispositivo moral que refuerza las demandas de la moral convencional (cf. Baier 1993, 446). Correspondientemente, la alabanza, como opuesta a la vergüenza, tiene la misma función de control. La imagen general sobre la moralidad que podemos derivar de su posición sobre la infidelidad es la de una moral centrada en el escarnio público y en la detección de fallas, tanto como en el “aplauzo” y la “alabanza” (cf. Baier 1993, 449-450). En otras palabras, parece que para Hume la moralidad consiste en una práctica esencialmente social, que contiene elementos de control públicos que van desde la vergüenza, la acusación pública, la mala fama, hasta la alabanza, la buena fama y el reconocimiento y que, característicamente, reúne aspectos reflexivos y sentimientos morales distintivos. Respecto de los sentimientos, Hume dice que los sentimientos que marcan a las acciones virtuosas, en el terreno de la moral, son la amabilidad y la belleza propias de las virtudes sociales. Tal es la conclusión del argumento donde prueba que la utilidad como criterio moral no podría surgir por un acto de convención sino, como vimos arriba, por las determinaciones naturales de nuestra mente que la hacen placentera. Con base en esto, Hume concluye entonces:

Por consiguiente, hay que conceder que las virtudes sociales tienen amabilidad y belleza naturales que desde un principio anteceden a todo precepto o educación, y que las recomienda a la estima de la humanidad ignorante, y atrae sus afectos. Y como la utilidad pública de estas virtudes es la principal circunstancia de la que derivan su mérito, se sigue que el fin que tienden a promover debe resultarnos agradable de algún modo o por consideraciones de autointerés o por miras y motivos más generosos (IPM SB 215).

Según este argumento, la moralidad surge por ciertas determinaciones básicas de nuestra naturaleza, relativas al honor, la vergüenza, la amabilidad, lo odioso, etc., las cuales son

es un ejemplo impactante de la crueldad del razonamiento kantiano (cf. Baier 1993, 446). Por contraste con las ideas de Hume, que también estudia Baier, yo estoy de acuerdo con su apreciación.

⁹⁸ He modificado la traducción de F. Duque, quien eligió ‘infamia’ en vez de ‘vergüenza’, para mantener la consistencia terminológica y resaltar el punto de Hume sobre la vergüenza, a la que considera una emoción normativa básica.

emociones reactivas, esto es, emociones que surgen en un ámbito social como *reacciones* o respuestas a situaciones donde intervienen varios agentes (el orgullo, por ejemplo, no es una emoción reactiva de este tipo, pues puede experimentarse sin necesidad de otros agentes). El caso de la infidelidad ilustra el modo en que la emoción de la vergüenza desempeña un complejo papel, como emoción moral, como dispositivo restrictivo de control social y, muy importante, como juicio valorativo con cierto contenido cognitivo, esto es, como un juicio empleado para presentar a un cierto acto como *vergonzoso* y, simultáneamente, evaluarlo negativamente (volveré un poco más adelante a considerar este tipo de juicio, en relación con la cuestión del naturalismo de Hume).

Este carácter reactivo o social de las emociones básicas mencionadas, que está en la génesis de las prácticas morales, se relaciona posteriormente con las virtudes sociales, las cuales comprenden acciones que contribuyen a la utilidad social o bienestar común, y son así aprovechadas por políticos (entre otras figuras públicas) para formular convenciones morales que contribuyen al bienestar social. Pero incluso estas convenciones, según Hume, son a su vez sancionadas por los sentimientos placenteros relativos a la amabilidad y belleza. De modo que las virtudes sociales, aunque producidas por actos de convención, se vinculan con características emotivas que actúan con eficacia normativa y motivadora. En un agente individual, estas características emotivas reúnen prácticamente todos los aspectos de la mente y dejan ver, en funcionamiento, el dinámico sistema moral de la motivación humeana, el cual incluye a una cierta capacidad reflexiva y la capacidad de cada agente de amoldar sus disposiciones hacia el punto de vista social. Esta es la plataforma psicológica que hace posible que adopte el punto de vista del observador ideal. Nuevamente, las emociones básicas que hemos considerado juegan aquí el papel principal. Al respecto, la fuerza motivadora y normativa principal es la fama, esto es, el aspecto vinculado con la vergüenza, si bien de una manera opuesta. La siguiente cita deja ver cómo operan estos diversos elementos en un complejo acto de motivación y de seguimiento de ciertos patrones o reglas sociales:

Otra fuente de nuestra constitución que proporciona una gran adición de fuerza al sentimiento moral es el deseo de fama, el cual gobierna con tal autoridad incontrolada en todas las mentes elevadas, y es a menudo el objetivo principal de todos sus propósitos y empresas. Mediante nuestro intento continuo y fervoroso de hacernos con un carácter, un nombre y una reputación en el mundo, frecuentemente pasamos revista a nuestro comportamiento y conducta, y consideramos cómo aparecen a los ojos de aquellos que se nos acercan y nos observan. Este continuo hábito de examinarnos a nosotros mismos reflexivamente, por decirlo así, mantiene vivos todos los sentimientos de lo correcto y lo erróneo, y engendra en las naturalezas nobles una cierta reverencia tanto por ellas mismas como por los demás; lo cual es el guardián más seguro de toda virtud. Las comodidades y placeres animales pierden gradualmente su valor, mientras que toda belleza interior

y gracia moral se adquieren con aplicación, y la mente logra toda perfección que pueda adornar o embellecer a una criatura racional (IPM 9.1 SB 276)

Desde el punto de vista de un moralista kantiano, éste es el terreno de la moral heterónoma. La imagen de un agente moral presente en este pensamiento no es la de un agente que decide aislado, reflexionando sólo sobre el mejor modo de cumplir con el deber moral, independientemente de la opinión de cualquier otro agente, aspirando a condiciones de la mayor autonomía posible. Al contrario, el agente moral humeano es un agente social, que ajusta constantemente su carácter de acuerdo con la buena fama, esto es, la opinión favorable que los demás tengan de él o ella. Su conducta se ve dirigida, o restringida, por consideraciones de la opinión que los demás tienen sobre su persona, así sea en el sentido de fomentar su buena fama o de evitar la vergüenza. Este agente interioriza los estándares sociales de lo apropiado y lo inapropiado y, a su vez, él mismo es un difusor de tales estándares, porque también es juez moral de la conducta de sus pares y contribuye con su juicio a su buena o mala fama. Ésta es, quizá, una moral de la maledicencia y del prejuicio (cf. T 3.2.12.4 SB 571), si bien orientadas a que cada persona desarrolle sus capacidades del mejor modo posible en un determinado ámbito social y, además, que contribuya a la utilidad y bienestar sociales.⁹⁹ Pero de este modo los agentes humeanos sólo afinan su sensibilidad moral, esto es, los sentimientos morales relativos a lo correcto y lo erróneo. Consistentemente con la teoría de Hume del juicio moral, los criterios morales relacionados con la buena fama y la vergüenza son básicamente ciertos sentimientos, esto es, ciertos estados mentales que no surgen de la deliberación racional, aunque no son ajenos a ella, pues en la consideración del objeto moral, la acción virtuosa, se pone de relieve «la flexibilidad de los sentimientos y las distintas modificaciones de tal objeto cuando lo considera nuestra mente, al evaluarlo y juzgarlo (cf. T 3.3.5.6 SB 617). Consideraré ahora el tercer asunto que le interesa a Hume sobre la relación entre la psicología individual y las determinaciones sociales.

⁹⁹ Estoy de acuerdo con Baier (cf. Baier 1993, 446) en que está moral parece más plausible que la alternativa kantiana, que tiene claros rasgos de crueldad, según vimos en la cita 97, p. 168. Por contraste con la interpretación de Baier, me parece que yo le doy más peso que ella a la utilidad social de esta moral de la maledicencia y los prejuicios. A favor de Hume, hay que señalar que la moral kantiana, en sus reflexiones más extremas sobre el respeto a la ley, no parece concede ningún peso a la utilidad social, según se ve, por ejemplo, en el siguiente pasaje de las MM: “Incluso si una sociedad civil fuera a disolverse por consentimiento de todos sus miembros (e.g., si la gente de una isla deshabitada decide separarse y dispersarse por el mundo), el último asesino en prisión primero tendría que ser ejecutado, de manera que así todos le han hecho lo que sus acciones merecen y la sangre culpable no se aferra a la gente por no haber insistido en su castigo” (MM, 6:333).

5.4. Cómo las respuestas de los agentes morales son consistentes entre sí: los elementos normativos de la teoría moral de Hume

La teoría moral de Hume, tal como la he presentado, incluye elementos normativos diversos. Siguiendo a Beauchamp (2008), voy a entender por norma a la regla o estándar con autoridad para juzgar o dirigir al razonamiento, la conducta o las creencias humanas. Así, una norma guía, ordena, requiere, restringe o elogia. Correspondientemente, el incumplimiento de una norma en una situación dada garantiza censura, crítica, condena, desaprobación o alguna valoración negativa relacionada. Las normas se aplican a los usos lingüísticos, al método científico, a la epistemología, a la conducta moral, a la ley, a la práctica legal, a las normas de urbanidad o de etiqueta, etc. Por su parte, ‘normativo’ es la forma adjetival de ‘norma’. Según Beauchamp, los criterios de Hume para formar, corregir y justificar inferencias y creencias son las fuentes de su normatividad. Las creencias que requieren de corrección son suscitadas por condiciones de información deficiente o insuficiente, prejuicios, supersticiones, exceso de confianza en los testimonios humanos, metafísica especulativa, metodología inadecuada, inferencias causales apresuradas y ciertas ficciones de la mente. Además, para Beauchamp el núcleo de la moralidad de Hume reside en ciertas virtudes de carácter que son universales y que establecen autoridad (cf. Beauchamp 2008, 493-494). En otras palabras, la fuente de la normatividad moral, para Hume, residiría en el desarrollo y mantenimiento de ciertas disposiciones para juzgar, pensar y actuar sobre cuestiones morales. Ahora bien, hemos visto que el sentimiento moral, que por su estructura mental podría ser una pasión indirecta, está constituido por ciertos elementos cognitivos que funcionan como juicios.

Estos juicios, de acuerdo con las ideas de Beauchamp, también funcionan como normas y son susceptibles de ser corregidos si se suscita alguna de las condiciones indicadas antes. Sin embargo, esta característica no es suficiente para sostener que la teoría moral de Hume es racionalista porque, como señala Beauchamp, para Hume la moralidad es contingente sobre la naturaleza humana, la cual por sí misma es la fuente de la moralidad universal (cf. Beauchamp 2008, 493). Es decir, las virtudes de carácter mencionadas antes están de alguna manera determinadas por la propia naturaleza humana y, según hemos visto, consisten en ciertas disposiciones naturales, que Hume llama virtudes naturales, las cuales «progresan» socialmente y, eventualmente, dan lugar a

otras disposiciones determinadas por convención (cf. antes §4.2.1).¹⁰⁰ En germen, esta idea de Hume contribuye a la explicación que quiero desarrollar en esta sección: cómo las respuestas de los agentes morales son consistentes entre sí. De acuerdo con Hume, la consistencia en nuestras respuestas morales proviene en primer lugar de las determinaciones naturales detrás de la moralidad. No podemos sino experimentar al mundo humano en términos morales y esta experiencia es potenciada por la internalización de ciertas normas –las correlativas con las virtudes artificiales. Hume logra articular de este modo las condiciones básicas de nuestra naturaleza moral con las condiciones sociales de la moral.¹⁰¹ Voy a complementar la explicación sobre el proceso social de interiorización de las normas que he desarrollado desde las secciones anteriores, siguiendo una distinción que, incluso en el planteamiento de Beauchamp, está presente en la teoría de la normatividad moral de Hume: por una parte, podemos describir los elementos normativos presentes en la teoría moral de Hume (como una especie de inventario de las normas morales) y, por otra parte, estudiar las condiciones de *corrección* de tales elementos normativos. Creo que al respecto, el recurso más importante y afín al pensamiento de Hume es la teoría del observador ideal. Ahora bien, del aspecto descriptivo mi interés para esta sección recae sobre las normas sociales establecidas por convención porque estrictamente las determinaciones naturales no constituyen normas y, según Hume, de ellas se derivan las normas sociales. De éstas consideraré las relativas a la justicia, la propiedad y el cumplimiento de las promesas. En consecuencia, la cuestión a la que me dirigiré puede plantearse ahora en términos de cómo, de acuerdo con Hume, nuestras respuestas morales son consistentes en relación con las normas de la justicia, de la propiedad y las promesas.

Si tales normas surgen por algún acto de acuerdo o convención, ¿en qué sentido tenemos que entender la consistencia en las respuestas morales relativas a ellas? El origen mismo, por acuerdo o convención, garantizaría la consistencia. Sin embargo, las ideas de Hume, una vez más, parecen más interesantes y plantean retos importantes de comprensión. Por una parte, él parece sostener algo más fuerte que un mero acuerdo cuando afirma que “[A]unque las reglas de justicia sean *artificiales*, no son arbitrarias” (T 3.2.1.19 SB 484), porque el contenido del acuerdo podría

¹⁰⁰ Las virtudes o determinaciones morales naturales comprenden la afabilidad, la beneficencia, la caridad, la generosidad, la clemencia, la moderación y la equidad, cf. (T 3.3.1.11 SB 578).

¹⁰¹ Hume reconoce, al menos en un pasaje (T 2.3.1.9 SB 402), que las diferencias entre seres humanos se deben a las diversas y necesarias condiciones sociales, toda vez que surgen de los “necesarios y uniformes principios de la naturaleza humana”.

ser arbitrario y la afirmación de Hume sugiere una cierta necesidad en las reglas de justicia. Por otra parte, Hume reconoce que hay una gran variación en sistemas normativos similares. Vimos antes que considera que a pesar de la naturaleza humana, las diferencias entre los seres humanos se deben a la sociedad (cf. T 2.3.1.9 SB 402). En la IPM, discutiendo el origen de las reglas de propiedad, reconoce que si no se tiene en cuenta que el fin de tales reglas “es el interés y la felicidad de la sociedad humana”, nada podría “parecer más caprichoso, antinatural e incluso supersticioso que todas o la mayoría de las leyes de la justicia y la propiedad”. Para ilustrar esta idea señala que mucha gente suele «mofarse» de supersticiones populares relativas a la “consideración especial por determinados tipos de carne, o días, lugares, posturas e indumentarias” (cf. IPM 3.2 SB198-199). Es decir, Hume reconoce que los sistemas normativos y las sociedades humanas varían mucho entre sí, en costumbres, creencias, normas particulares, etc. y que desde un punto de vista racional estas variaciones parecen arbitrarias. En efecto, a primera vista no parece haber ninguna razón inteligible para que ciertas culturas, por ejemplo, se abstengan de comer carne de cerdo o de vaca, o para que sea el anfitrión quien sale primero de su casa y deja en libertad a sus invitados de salir cuando quieran o que sean éstos los primeros en salir y el anfitrión el último, etc. (cf. IPM 199 y ss., así como 337 y ss., para muchos otros ejemplos de variaciones de este tipo). Con todo, la posición de Hume parece ser que estas variaciones son superficiales y que hay *niveles* donde las normas morales son uniformes para toda la humanidad.

Para empezar, un examen atento muestra que todas ellas contribuyen, a su modo particular, «al intereses y felicidad de la sociedad humana». Hume parece decir que a pesar de las variaciones que podamos encontrar en los diversos sistemas normativos, su finalidad es el bien común y la felicidad de la sociedad por medio de la práctica de la justicia, el respeto a la propiedad privada y, en general, el cumplimiento de las promesas. De éstas, el respeto a la propiedad privada es para Hume la regla que puede especificarse más y comprende tres reglas más particulares relativas a la estabilidad de la posesión, la transmisión mediante un procedimiento acordado colectivamente y el cumplimiento de las promesas, la cual garantiza el cumplimiento de las dos anteriores. Aun cuando pueda haber algunas variaciones en cómo las diferentes sociedades entienden y ejecutan estas reglas generales, no parece que pueda haber, según él, ninguna sociedad humana sin ellas:

...aunque a los hombres les sea posible mantener una pequeña e inculta sociedad sin gobierno, será imposible que mantuvieran una sociedad, del tipo que fuere, sin justicia y sin la observancia de las tres leyes fundamentales de la estabilidad de posesión, su transmisión por consentimiento y el cumplimiento de promesas (T 3.2.8.3 SB 541).

Para Hume, de acuerdo con esta idea, las reglas de la moral son un componente intrínseco a las sociedades humanas y su respeto es «el origen de la obediencia civil». Discutiendo el origen de las sociedades políticas y esta misma necesidad que tienen de las reglas de justicia, afirma Hume este mismo pensamiento: “La naturaleza humana no puede subsistir de ningún modo sin la asociación de individuos; y esta asociación nunca podría tener lugar si no se respetan las leyes de la equidad y la justicia” (cf. IPM 4.3 SB 20). Así, aunque las formas particulares de la moralidad en una sociedad dada podrían ser contingentes en el sentido de que podrían ser diferentes a lo que de hecho son (tal vez se consideraría virtuoso o vicioso comer carne de tal o cual animal en tal o cual ocasión, etc.), la institución misma de la moralidad no es contingente, sino necesaria a la propia sociedad. En efecto, no parece que podamos concebir sociedad alguna sin normas morales aceptadas, de alguna manera, por los miembros integrantes de la misma. Esta es una idea que no suele reconocerse como propia del pensamiento moral de Hume. Además, la relevancia de estas ideas para explicar la motivación moral es patente en el pensamiento de Hume, lo que muestra que para él hay un continuo respecto de las motivaciones despertadas de manera natural y las que responden a las reglas sociales por convención.

Así, aun cuando el respeto a la regla de justicia es fundamental para el mantenimiento y preservación de la sociedad, Hume además afirma que afecta al sentimiento moral, la principal fuente de la motivación humana:

La necesidad de la justicia para el mantenimiento de la sociedad es el único fundamento de esa virtud; y puesto que ninguna excelencia moral es más altamente estimada, podemos concluir que esta circunstancia de la utilidad tiene, en general, la energía más fuerte y el dominio más completo sobre nuestros sentimientos (IP 3 SB 203-204)

Esta poderosa fuerza que influye sobre nuestros sentimientos es, con todo, «artificial» en el sentido de que el respeto a la justicia y las ideas relativas a ella no tienen su origen en las determinaciones naturales. Al mismo tiempo, es la energía motivadora que conduce a los agentes a respetar *consistentemente* las obligaciones y deberes sociales de los que son receptores. Esta consistencia, en la teoría de Hume, es el fundamento mismo del «partido de la humanidad» y se opone a las tendencias egoístas de la naturaleza humana al mismo tiempo que fortalece progresivamente la consideración social y la «benevolencia»:

¿Qué tiene de extraño, entonces, el que se encuentre que los sentimientos morales tienen tanta influencia en la vida, aunque surjan de principios que a primera vista pueden parecer un tanto egoístas y débiles? Pero tenemos que observar que estos principios son sociables y universales. De alguna manera forman el partido de la humanidad contra el vicio o el desorden, su enemigo común; y, según la preocupación benévola por los demás se difunde en mayor o menor medida entre todos

los hombres, y es la misma en todos ellos, aparece con más frecuencia en el discurso, es apreciada en la sociedad y en la conversación, y la censura y la aprobación consecuentes a ella se despiertan de este modo de ese letargo en el que probablemente permanecen en la naturaleza solitaria y no cultivada (IPM 9 SB 275-276).

La consistencia en las respuestas morales de los agentes es reforzada socialmente, mediante la alabanza de las prácticas de benevolencia y consideración por lo demás. Como hemos visto, ésta es una moral de la forma de un observador ideal, esto es, que explica a las acciones y pensamientos morales como una respuesta a un determinado punto de vista ideal. La consistencia en las respuestas morales se explica porque los agentes comparten este punto de vista ideal, el punto de vista de la humanidad, el cual les señala el estándar normativo apropiado. Ésta es la fuente para que los agentes generen motivos apropiados para cumplir con su deber y obligación. Sin embargo, la fuerza de este estándar normativo es débil, particularmente en relación con los deberes y obligaciones.

En la concepción de Hume, la función principal de los deberes y obligaciones civiles o sociales, y por extensión la del gobierno que los hace cumplir, es preservar a los deberes naturales, los correlativos con las virtudes naturales: “Nuestros deberes *civiles* están, pues, conectados con los *naturales* en cuanto que los primeros fueron inventados fundamentalmente para salvaguardar los segundos, y en cuanto que el objeto principal del gobierno es obligar a los hombres a observar las leyes naturales” (T 3.2.8.5 SB 544). A pesar de esto, desde el punto de vista de los agentes particulares, estos deberes y obligaciones son percibidos como remotos, especialmente los relativos a la necesidad misma de la justicia para la sociedad. Es posible que los agentes individuales sientan claramente las restricciones particulares, como llegar a cierta hora, atender ciertas tareas, etc., y no sientan tan claramente –o de ninguna manera en absoluto– cómo contribuyen así al bienestar colectivo y a la felicidad de la sociedad. Éste es un rasgo de la psicología humana que Hume explica así: “Si examinamos el curso ordinario de las acciones humanas, hallaremos que la mente no se refrena en base a reglas generales y universales [como los deberes y obligaciones más fundamentales del acuerdo social], sino que su actuación está determinada en la mayoría de los casos por los motivos e inclinaciones del momento” (T 3.2.6.9 SB 531). Quizá Hume explica así la negligencia al cumplir los deberes y obligaciones. Aun cuando para un agente particular las demandas que éstos le imponen resulten inmediatas y explícitas, muy posiblemente tales demandas estarán en desventaja cuando entren en conflicto con deseos o apetitos relativos a placeres personales. La placentera perspectiva de tomarse un café o una

cerveza podría ser más fuerte que la de volver al trabajo o revisar un expediente más, etc. Ahora bien, de acuerdo con la teoría de la motivación de Hume, este conflicto sólo puede resolverse, en el sentido que sea, si se entiende como un conflicto entre pasiones o deseos (i. e., entre estados no-cognitivos). Como hemos visto antes, para Hume una distinción moral sólo puede entenderse en términos de un cierto sentimiento o pasión hacia un cierto objeto. Por ejemplo, vimos que para él una acción «sólo puede ser considerada buena o mala si una pasión o motivo nos impele a ella» (cf. T 3.2.2.18 SB 486 y T 3.2.6.9 SB 532; v. p. 143). Esta idea cubre también a los deberes y obligaciones.

Los agentes humeanos cumplen las demandas de sus deberes y obligaciones sólo si son impelidos a ello por una pasión. Hume concibe esto como una ampliación de los deberes naturales, según hemos visto. Según él, los deberes sociales se entienden como una ampliación natural de los deberes filiales:

Un hombre ama naturalmente más a sus hijos que a sus sobrinos, a éstos más que a sus primos y a estos últimos más que los extraños, siempre que todas las demás circunstancias sean iguales. De aquí es de donde surgen nuestras reglas comunes del deber, prefiriendo unos a otros. Nuestro sentido del deber sigue en todo momento el curso común y natural de nuestras pasiones (T 3.2.2.18 SB 483-484).

«El sentido del deber» y la constricción correspondiente que un agente experimenta por un deber del que es receptor, de acuerdo con esta idea, consiste en una cierta pasión o estado no-cognitivo que impele al agente a actuar. Como vemos, por contraste con las determinaciones normativas naturales mencionadas ésta es una fuerza débil. Sin embargo, los factores sociales y la propia experiencia social del agente la fortalecen y logran darle los poderes normativos que la vida social requiere. Desde la perspectiva del agente individual, según Fate Norton, esto lo determina a aceptar a las normas sociales como si fueran normas *necesarias*, bajo una forma de pensamiento similar a la que el mismo agente acepta la necesidad causal en el mundo natural:

...nuestra idea de deber es la consecuencia de un proceso de experiencia que es estructuralmente similar al proceso que da lugar a la idea de conexión necesaria. [...] Las observaciones de Hume acerca del deber sugieren [Hume no parece desarrollar así este punto] que una impresión similar de las expectativas subyace a nuestra idea de la obligación moral, así como de la culpa, y al menos algunas formas de aprobación (Fate Norton 2003, 170).

Los marcos sociales que encuadran a la acción individual, así como la experiencia de los agentes individuales, determinan que el agente conciba que los requisitos normativos se ejerzan sobre él con una fuerza irresistible o necesaria, de un modo semejante a cómo considera la causalidad en el mundo natural como un factor necesario. En efecto, al menos en un lugar, cuando describe a la

relación de propiedad, Hume dice que es “una especie de causa y efecto” (cf. T 3.2.3.7 SB 506), por la capacidad que tienen los agentes para afectar a los bienes que poseen, pero también es claro que describe una regla o norma, esto es, la regla de propiedad. Ahora bien, a pesar de que esta última analogía no la hace Hume, sus ideas sobre la formación de creencias y sobre el proceso de interiorización de normas hace plausible esta explicación. Según esta analogía, en la estructura de este proceso de formación de creencias causales o de motivos para actuar no participa ninguna facultad cognitiva. Los elementos que sí están presentes son la experiencia, el hábito o la costumbre. Estos elementos, combinados de una u otra manera, parecen ser los materiales básicos de las normas, en el ámbito epistemológico o en el moral, según esta explicación general de Hume. Yo considero que la simpleza y alcance general de esta explicación es un rasgo más de la consistencia del pensamiento de Hume. Con un mínimo de elementos logra explicar diversos fenómenos relativos a la psicología humana y su operación en ámbitos muy distintos, en este caso, el ámbito de la formación de creencias sobre cuestiones de hecho y el ámbito de la formación de motivos para actuar de acuerdo con el deber y la obligación social. Hume describe a este rasgo con la expresión «la flexibilidad de sentimientos» y con él se refiere precisamente a la semejanza que hay en los sentimientos de aprobación hacia objetos muy distintos entre sí, como los mencionados antes:

Puede que no esté ahora de más el poner de relieve la flexibilidad de nuestros sentimientos y las distintas modificaciones que tan fácilmente reciben de los objetos con que están conectados. Todos los sentimientos de aprobación que acompañan a cualquier especie particular de objetos tienen gran semejanza entre sí, a pesar de derivarse de fuentes distintas (T 3.3.6.6 SB 617).

Entre los objetos que Hume menciona están la belleza de los objetos sensibles, las acciones y el carácter de los seres humanos, así como la utilidad de otros, como una casa. Ante estos objetos, experimentamos de manera similar sentimientos de aprobación y placer ante ellos, aunque exista diversidad respecto de las causas que originan tales sentimientos. Hume confina al terreno del misterio este aspecto relativo a las causas, me parece porque no es claro si los sentimientos de aprobación y placer, o sus correlatos de desaprobación y desplacer, se originan en la estructura de la mente humana o si algunas propiedades de los objetos en cuestión, además de su utilidad, intervienen también en su origen. A pesar de esto, es un hecho, para Hume, que experimentamos consistentemente los sentimientos de aprobación y desaprobación en nuestras pasiones y sentimientos, tal como presenta este pensamiento: “Existe algo realmente inexplicable en esta variación en nuestra afectividad; sin embargo, tenemos experiencia de ello en todas nuestras

pasiones y sentimientos” (T 3.3.6.6 SB 617). Por lo demás, esta flexibilidad de sentimientos, que caracteriza en buena medida a la estructura valorativa de la mente humana, no sería posible sin la simpatía.

Como vimos en el capítulo dedicado a las pasiones, la simpatía no es ella misma una pasión, pero sí una facultad o capacidad de la mente que hace posible que se despierten las pasiones (v. §2.2.3). En este capítulo he desarrollado la idea de que el sentimiento moral tiene la estructura de una pasión indirecta, de modo que es consistente con estas ideas. Ahora examinaré cómo la simpatía ejerce sobre el sentimiento moral una influencia análoga a la que ejerce sobre las pasiones. En el argumento de Hume, dicha influencia se refiere a las virtudes sociales o artificiales, porque desde su punto de vista las virtudes naturales no parecen requerir de ningún auxilio para ejercer su influencia en los agentes a favor de otros agentes. La simpatía, entonces, actúa sobre el sentimiento moral para dirigir a la mente hacia otros agentes:

...es manifiesto que la simpatía es un principio muy poderoso en la naturaleza humana, que tiene gran influencia en nuestro sentido de la belleza y que origina el sentimiento moral en todas las virtudes artificiales. Partiendo de estos puntos, cabe suponer que la simpatía origine también muchas de las restantes virtudes y que sea la tendencia al bien de la humanidad lo que haga merecedora de nuestra aprobación a una cualidad mental (T 3.3.1.10 SB 577-578, énfasis de Hume).

En el *Tratado*, según se desprende de estas ideas, la simpatía es la fuerza o capacidad de la mente humana para despertar el sentimiento de belleza y la aprobación hacia los bienes relativos a las acciones humanas, así como el placer que los agentes humanos experimentan ante ellos. Así, la simpatía es el elemento decisivo detrás de la sociabilidad de las pasiones. Al parecer, las pasiones sin la simpatía son inertes, pero gracias a ella se vuelven sociales y dirigidas hacia otros agentes (cf. McIntyre 212-213). En otro aspecto de las virtudes artificiales, por ejemplo el relativo a la justicia o a la propiedad privada, la simpatía, en tanto que una capacidad de la imaginación (cf. Stroud 285), refuerza a la percepción que los agentes tienen de las reglas generales, las cuales comprenden a los deberes y obligaciones particulares. Esta consideración permite, eventualmente, que las facultades cognitivas de la mente, como lo es en algún grado la imaginación, pero también la reflexión, influyan sobre los agentes en la formación de los motivos para cumplir con el deber y la obligación moral. Hume explica así esta relación:

La imaginación tiene un grupo de pasiones de que dependen en gran medida nuestros sentimientos de belleza. Estas pasiones son movidas por grados de vivacidad y vigor inferiores a la creencia y, por tanto, independientes de la existencia real de sus objetos. Cuando un carácter es apropiado por todos conceptos para resultar provechoso a la sociedad, la imaginación pasa fácilmente de la causa al efecto sin preocuparse de que todavía falten algunas circunstancias necesarias para que la causa

sea completa. Las reglas generales originan una especie de probabilidad que a veces influye sobre el juicio y siempre sobre la imaginación (T 3.3.1.20 SB 585).

Esto es suficiente para que admitamos que la reflexión –la razón–, impulsada por la imaginación y este «cálculo de probabilidades» que le prepara, pueda intervenir en el proceso de formación de motivos y pasiones detrás de las pasiones artificiales y los deberes sociales. A diferencia del modo en que vimos cómo la razón puede influir sobre las pasiones por medio del «juicio que acompaña» a aquellas (v. §2.2.1), en el caso del deber y la obligación la reflexión puede ser, directamente, un factor que influya sobre el sentimiento moral. Según Hume, la idea es que “[L]a reflexión sobre las tendencias de nuestra acción tiene con mucho la mayor influencia, y determina las grandes líneas de nuestro deber” (T 3.3.1.27 SB 590). En otras palabras, la imaginación y la razón, en el ámbito de la moral, suministran elementos para que los agentes puedan reflexionar sobre sus acciones, las expectativas que recaen sobre ellos, los juicios de otros agentes, sus pares, y sobre los deberes y obligaciones de que son receptores. Estos elementos fortalecen al sentimiento y a la pasión moral y determinan, en último lugar, que el agente encuentre los motivos que lo impelen a actuar de acuerdo con las demandas de la moral.

6. Consideraciones finales: Valoración del naturalismo moral de Hume y el papel que da a la belleza y a la utilidad. Defensa de la interpretación del observador ideal

Al principio de la sección 5.1., cité un pasaje de la “Introducción” del *Tratado* que puede ser leído como la *agenda* del naturalismo de Hume. La afirmación central de este naturalismo es que «todas las ciencias, incluso las matemáticas, la filosofía natural –o ciencia de la naturaleza–, la religión, dependen de la ciencia del hombre» (cf. T I 4 SB xv). De acuerdo con esta agenda, la investigación de la naturaleza del *hombre* –de los seres humanos– nos dará la clave para comprender a «todas las demás ciencias». Hume se propone desarrollar esta investigación en el *Tratado* y lo hace, me parece, elaborando una *metodología* que consiste en determinar el origen de nuestras ideas en la mente a partir de las impresiones sensibles y las de reflexión (T 1.1.2. SB 7). Después de todo, las ciencias humanas y aún la religión se componen de ideas, de modo que la afirmación de que dependen de la naturaleza humana puede entenderse como que dependen de la manera en que se forman las ideas que las constituyen. Así, su teoría de las ideas es la base de todo su sistema filosófico, incluyendo sus explicaciones sobre la moralidad. Ahora bien, sabemos que las ideas y creencias se forman por tres principios generales, que explican las regularidades de nuestro pensamiento en todos los órdenes y que juntos explican el funcionamiento de la facultad de la imaginación: la semejanza, la contigüidad en tiempo y espacio, y la causa y efecto (cf. T 1.1.4.1 SB 10-11). En resumen, todas las ciencias deben sus creencias a estos principios generales. ¿Qué tan plausible es esta explicación?

Creo que la primera dificultad para responder esta pregunta consiste en determinar cómo puede responderse satisfactoriamente. Stroud, por ejemplo, se ha propuesto valorar el naturalismo de Hume de manera muy apegada a la idea de que todas las ciencias dependen de la naturaleza humana. En consecuencia, para él la tarea de valorar este naturalismo consiste en responder «qué aportaciones hace a nuestra comprensión de la naturaleza humana y cuánto puede ganarse filosóficamente siguiendo sus estudios al respecto» (cf. Stroud, 324). Pero según él cometeríamos un error si juzgamos a Hume como “un científico social o humano contemporáneo”. La cuestión para Stroud no es hasta qué punto las ideas de Hume son vigentes según los estándares de las ciencias contemporáneas, sino cuánto nos enseñan *filosóficamente* sobre la naturaleza humana. Pero aún esto lo admite con reservas, pues reconoce que las explicaciones de Hume “son a menudo

falsas en sus detalles, simplistas en su concepción, y que apenas *explican* el fenómeno particular que *explican*". De manera que desplaza el foco de la respuesta "a los problemas planteados, o los temas expuestos, incluso tanto como la posibilidad del tipo de investigación o del modo de comprensión que propone" (cf. Stroud, 325). Hume, según esta visión, es más importante por las preguntas que planteó, por los temas en que se interesó y por explorar una cierta metodología para estudiarlos. Desde luego, admito sin reservas este dictamen de Stroud, pero creo que sobre la motivación y la moralidad podemos ir un poco más lejos y considerar con atención las explicaciones de Hume.

En esta línea, por ejemplo, T. Beauchamp sostiene que podemos encontrar en el naturalismo de Hume un lugar para los objetos morales, de acuerdo con el siguiente razonamiento:

...los enunciados acerca de los sentimientos son enunciados fácticos, y no simplemente enunciados acerca de los objetos a los que se dirigen tales enunciados. Los hechos acerca de la moral se derivan de hechos acerca de los sentimientos en tanto que estos sentimientos están dirigidos a objetos (siendo éstos cualidades mentales). Interpretado así, Hume parece ser un naturalista en teoría moral (cf. Beauchamp 2008, 500).

De acuerdo con Beauchamp, los sentimientos morales son hechos que podemos estudiar directamente, es decir, que el estado de aprobación o desaprobación que un agente experimenta en alguna circunstancia moral son hechos *objetivos* que pueden ser caracterizados por algún procedimiento empírico. Con todo, Beauchamp reconoce que este planteamiento es compatible con una concepción tanto objetiva como subjetiva de los juicios morales (en un sentido amplio que parece comprender tanto valoraciones como normas), aunque él sostiene que una interpretación correcta de Hume es compatible con una concepción objetiva de los juicios, particularmente los relativos a la virtud (cf. Beauchamp 2008, 500).¹⁰² Beauchamp también indica que hay interpretaciones más "robustas" de la objetividad de los juicios morales humeanos – consideré muy brevemente la interpretación de Fate Norton al respecto (v. arriba §4.2, p. 127 y ss). En este sentido, el pensamiento moral de Hume se ajusta bien a los criterios con los que solemos identificar al naturalismo moral hoy en día. Por lo menos ilustra la tesis de que hay hechos morales naturales y objetivos que, en principio, podrían ser estudiados por las ciencias empíricas (cf. Lenman 2014), esto es, los sentimientos morales.

¹⁰² Beauchamp no duda de que se trata de propiedades objetivas fuertes. Comentando T 3.3.1.19 SB 584 explica que la percepción de la regularidad de un carácter requiere del reconocimiento, aunque matizado por los sentimientos, de ciertas propiedades objetivas: "Regularidades objetivas (i.e. relata constantemente conjuntados) subyacen a los juicios de conexión causal; estos juicios causales adornan [*gild*] las regularidades mediante la proyección de los sentimientos de conexión necesaria" (cf. Beauchamp 2008, 499-500).

Consideraciones similares hace A. Baier, pero como una reacción contra las interpretaciones emotivistas de Hume de mediados del S. XX, las cuales explicaron su pensamiento moral meramente como subjetivista (p.e., Kenny 1963; Alston 1968; Deigh 1994). Según la interpretación de Baier, los juicios morales consisten en un especial punto de vista sobre aquellos rasgos de carácter *reales* que consideramos virtuosos, como «la serenidad o la jovialidad». Su interpretación presupone el punto de vista general o imparcial, a la manera de un juez: “...estas cualidades tienen que estar en algunas personas, a fin de que el juez moral les dé la bienvenida y los apruebe, y así las considere virtuosas”. Agrega que las “máximas indudables” de Hume garantizan alguna forma de realismo para su “epistemología moral” (cf. Baier 1991, 194). El punto que distingue a la posición objetivista de Beauchamp de la de Baier es el asunto de a *qué* se refiere el sentimiento moral. Ambos aceptan que el sentimiento moral es acerca de algo, pero mientras que para Beauchamp este algo es una cualidad mental, para Baier es alguna característica real de las personas que son objeto de los juicios morales. Creo que parte de las dificultades para ofrecer una correcta interpretación del pensamiento de Hume sobre el tema de los juicios morales es que, según ha mostrado Stroud, en sus textos hay pasajes que son compatibles con distintas concepciones de ellos (cf. Stroud 261 y ss).

En efecto, Hume sostiene en ocasiones que el objeto de las impresiones correspondientes al vicio o la virtud podrían ser meros estados subjetivos sin relación alguna con objetos externos a la mente (cf. T 3.1.2.3 SB 471; IPM SB 173); en otros, que son objetos propios del mundo *natural*, es decir, existentes independientemente de que los agentes los perciban de alguna manera (T 3.3.1.19 SB 584; IPM, 289; cf. Norton 1982 y 1985; Baier 1991, 193-5)¹⁰³ y, por último, objetos

¹⁰³ Ciertos pasajes del *Tratado* justifican esta posición, por ejemplo, en T 3.1.2.1 SB 470 Hume afirma que es “preciso reconocer, en efecto, que la impresión surgida de la virtud es algo agradable, y que la procedente del vicio es desagradable.” Más adelante parece ilustrar esta idea con el reconocimiento de ciertas propiedades de una persona que la hacen virtuosa:

Cuando alguien es de un carácter tal que su tendencia beneficia a la sociedad, juzgamos que esa persona es virtuosa, y nos produce placer contemplar su carácter, aunque no le sea posible hacerse patente por circunstancias particulares que incapaciten a esa persona para prestar servicios a sus amigos y país. Aun en harapos, la virtud sigue siendo virtud, y el amor que su contemplación origina sigue acompañando a un hombre en la cárcel o en el destierro, aun cuando la virtud misma no pueda ya ejercitarse y que esté perdida para el mundo (T 3.3.1.19 SB 584)

Vemos aquí la idea de que la virtud es un rasgo constante y regular de la persona, pero en este caso parece que se trata de una propiedad de la persona, la cual produce en quien la contempla placer, incluso si la persona virtuosa ha sido privada de la oportunidad de actuar. Además, Hume presenta aquí también un elemento importante de su concepción de la virtud: la utilidad (volveré a este elemento más adelante). De acuerdo con la idea, podríamos evaluar la *utilidad* del carácter o de la persona, por ejemplo, si ofrece «algún servicio a su nación o amigos». La propiedad de ser útil pertenecería de tal modo a la persona que incluso si no tiene oportunidad de ejercerla aun así

que de alguna manera existen por la participación de propiedades mentales y propiedades objetivas, una posición que suele identificarse como proyectivista o inter-subjetiva (cf. T 3.1.1.17-26 SB 463-469; DEG; IPM Ap. 1.21; Stroud, 184-185).

No pretendo defender ninguna de estas posiciones como la interpretación correcta de Hume. En todo caso, comparto la idea de la propia Baier de que no parece correcto fijar al pensamiento de Hume como ‘realista’, ‘no-realista’, ‘quasi-realista’ o ‘proyectivista’. Para hablar de las ideas de Hume sobre la evaluación moral podría ser más útil prescindir de estas categorías (cf. Baier 1991, 194). Es cierto, como sugiere su propia interpretación objetivista presentada arriba, que ella misma hace alguna concesión a «los gustos filosóficos actuales», pero creo que incluso podríamos prescindir de hacer concesiones como esa –que parecen, quizá, servir más a entender el pensamiento moral actual que el de Hume.

Considerando la relación que hay entre los juicios morales y las prácticas morales, en este capítulo me propongo responder la pregunta planteada antes: ¿qué tan plausible es la explicación de Hume que descansa en los principios generales de la formación de ideas? En este sentido, pretendo valorar su naturalismo moral considerando aspectos relativos a los usos y condiciones de aplicación de los juicios morales (entendidos como valoraciones morales y normas morales). No pretendo detenerme más en el estudio de los posibles objetos a los que se refieren. Es decir, intento hacer menos meta-ética de la teoría de los juicios morales de Hume y dirigir mi atención en la dirección opuesta –por decirlo así-, la que nos aleja de la teoría sobre la ética y nos acerca a las prácticas morales. Esto coincide con otro aspecto observación de Stroud sobre el valor del naturalismo de Hume: “Lo que atrae la atención es la idea misma de una investigación y una explicación empíricas totalmente exhaustivas de por qué los seres humanos son como son y por qué piensan, sienten y se comportan como lo hacen” (Stroud, 325). Creo que para responder esta cuestión es necesario examinar cómo se relacionan elementos como la utilidad, la benevolencia y

experimentaríamos el placer debido a la contemplación de la virtud. En la IPM Hume incluso ‘define’ a la virtud como si fuera una propiedad objetiva: “...la moralidad se determina mediante el sentimiento. Define a la virtud como *cualquier acción o cualidad mental que ofrece al espectador el sentimiento placentero de aprobación*; y al vicio como lo contrario” (IPM Ap. 1 SB 289). Al respecto, Stroud observa que en cualquier sentido que se tome, esta definición implica que la virtud y el vicio son rasgos objetivos de una acción o de una cualidad personal, los cuales provocan un sentimiento distintivo al observarlos o contemplarlos (cf. Stroud, 266). Stroud continúa explicando que si esta definición fuera correcta entonces por medio de la observación y el razonamiento podríamos descubrir qué acciones y caracteres son virtuosos y formular, correspondientemente, juicios morales basados solos en la observación y en el razonamiento. Stroud señala que esto es precisamente lo que Hume niega explícitamente (cf. Stroud, 266-267). En efecto, esta interpretación objetivista de las ideas de Hume parece ser contraria a las tendencias generales de su pensamiento.

la belleza, característicos de las valoraciones morales de los agentes humanos, con las prácticas morales ordinarias y con los elementos naturales básicos, esto es, las pasiones más básicas (amor, odio, orgullo y humildad). El propio Hume, como he presentado en el capítulo anterior, sostiene que el sentimiento moral *requiere* de factores vinculados a las determinaciones naturales básicas y a las condiciones sociales. Examinaré estos aspectos en las siguientes secciones.

Mi conclusión final es que todos estos elementos convergen bien en una interpretación general del pensamiento de Hume que de un lugar central a la interpretación del observador ideal. Según sostengo, el punto de vista social opera en cada uno de nosotros como una especie de plataforma psicológica que sostiene la vida mental de cada individuo agremiado a la sociedad. En otras palabras, nuestra forma de sentir, de pensar, de percibir es una construcción social, no racional, que opera en cada uno de nosotros y que articula todos los elementos mencionados. En buena medida, el punto de vista del observador ideal coincide con el punto de vista social y los criterios sociales.

6.1. La utilidad, las reglas morales y el sentimiento moral

Según lo dicho en el capítulo anterior, el sentimiento de aprobación moral no se origina en la utilidad sino en el placer proveniente del amor o en el desagrado proveniente del odio, los cuales son despertados por simpatía hacia otros agentes (cf. T 3.3.3.4 SB 604). También estudiamos, en el capítulo anterior (v. §5.2), que el conjunto de las reglas por el cual los agremiados de una sociedad regulan su conducta surge por un acto de convención, “un sentido general de interés común” (cf. T 3.2.2.10 SB 490). Según vimos, para Hume la fuerza restrictiva de las promesas, o su obligación, yace en factores sociales y en la utilidad que su cumplimiento trae para la sociedad. Para él la función social de las constricciones impuestas a los individuos por las reglas de la propiedad, la justicia y la promesa es contener «el curso natural de nuestras pasiones e inclinaciones naturales» y el egoísmo natural de los agentes humanos. De este modo, refuerzan las tendencias benévolas naturales de los agentes humanos y favorecen la realización de acciones «en beneficio de los demás de un modo desinteresado» (cf. T 3.2.5.8 SB 519-520). Nuevamente, el efecto sobre las pasiones es mediante el placer o el dolor que provocan en los agentes tales reglas y restricciones, esto es, mediante el único medio para motivar una acción o un pensamiento, según las ideas de Hume sobre la motivación. Presenté, además, que estos sentimientos de placer

o dolor son reforzados o expresados mediante dispositivos de control moral como la vergüenza, la alabanza o la fama (cf. sección 5.3).

El interés público es promovido por «los moralistas y los políticos», que tratan de «corregir la ingratitud y el egoísmo de los hombres», aun cuando sus poderes para influir al respecto son limitados y a lo más que pueden “pretender es dar una nueva dirección a esas pasiones naturales, enseñándonos que nos es posible satisfacer mejor nuestros apetitos de un modo oblicuo y artificial que siguiendo sus precipitados e impetuosos movimientos” (cf. T 3.2.5.9 SB 521). De este modo, Hume muestra que las determinaciones básicas de la naturaleza humana se pueden satisfacer de manera que puedan traer más provecho para todos. Wiggins considera que ésta es la característica más importante de la explicación genealógica de Hume sobre la moral:

Al final, para cualquiera que sienta y piense a través de todo lo que involucra, la virtud moral tiene que aparecer como el más fino florecimiento posible de las naturalezas originales y adquiridas del hombre –el adorno perfecto o embellecimiento del hombre considerado como un ser natural-cum-razonable (Wiggins 2006, 35).

Wiggins cierra este pensamiento remitiéndonos a IPM 9.1 SB 276, que hemos considerado antes (v. p. 170), donde Hume enfatiza que «los principios que dirigen a los sentimientos morales son sociales y universales; forman, de cierto modo, el *partido* de la humanidad en contra del vicio o desorden, su enemigo común». La moral, según esta explicación, es una derivación de las determinaciones naturales básicas de los agentes humanos, pero sus regulaciones y normas, sujetas a diversas contingencias culturales, históricas, etc., podrían expresar quizá los rasgos más distintivos de la naturaleza humana. Con todo, estas normas, según Hume, tienen que estar respaldadas por las determinantes naturales de la psicología humana.

En efecto, Hume afirma que no habría forma de que los agentes humanos actuaran de acuerdo con las reglas de la moral si no poseyeran previamente un motivo *distinto* al moral para actuar de acuerdo con ellas (cf. T 3.2.1.7 SB 479; Fate Norton 2000, 183). En este caso, la obligación de cumplir las promesas es un motivo interiorizado a favor del interés público y ensañado de maneras relativamente eficaces por diversos tipos de maestros y mediante diversos tipos de recursos. Parece, en todo caso, que estos recursos apelan a la experiencia, el hábito y la costumbre, y ocasionalmente a la reflexión. Los motivos morales, en este sentido, son meramente la expresión de convenciones, que bien podrían ser tácitas, a favor del interés común, y compartidos consistentemente entre todos los agremiados al pacto social, de acuerdo con lo que dice Hume en:

No hay que pensar que el conocimiento necesario para que los hombres se den cuenta del interés que tiene la institución y la observación de las promesas sea algo superior a la capacidad de la naturaleza humana, ni siquiera aunque sea ésta salvaje e inculta. Cualquiera persona las descubre al más leve contacto con la sociedad. Y como cada individuo percibe el mismo sentimiento de interés en sus semejantes, cumple inmediatamente su parte en el trato que haya efectuado porque está seguro de que los demás no querrán dejar de cumplir la suya. Todos ellos se conciertan en un esquema de acciones calculado para el beneficio común, esquema que se conviene será verdadero según la palabra acordada, y no se necesita de más, para formar este concierto o convención, que el que todos y cada uno de los individuos tengan un sentimiento de interés en el fiel cumplimiento de los compromisos que expresan ese sentimiento a los demás miembros de la sociedad, lo que produce que ese interés actúe sobre ellos; el interés resulta así la primera obligación para el cumplimiento de las promesas (T 3.2.5.11 SB 522-523).

La posición de Hume parece ser que, dentro del orden social, cuando los agentes son impelidos a actuar de acuerdo por las reglas morales casi sin notarlo, los motivos de todos tienden a converger en un “esquema de acciones calculado para el beneficio común”. Pero este cálculo no es realizado reflexivamente, sino que los agentes se *sintonizan* con él por medio de sentimientos placenteros o dolorosos, peculiarmente morales. Esta sintonía está ya presente en las etapas más «rudimentarias e incultas» de la sociedad, es decir, que no se requiere que se trate de acuerdos explícitos ni elaborados de manera sofisticada como podría ser una ley escrita y mediante instituciones sociales. De hecho, las instituciones sociales surgen precisamente de estas convenciones primarias dirigidas por el sentimiento, la experiencia y el hábito. Otra idea es que «la fuente de la normatividad», por usar una expresión conocida, es el propio interés social, el cual, nos dice Hume, es “así la primera obligación para el cumplimiento de las promesas”. En otras palabras, el sentimiento moral, en su manifestación social más básica, hace surgir en los agentes sentimientos favorables a las reglas morales, de modo que éstos experimentan con agrado cumplir con ellas, quizá adivinando de alguna manera su utilidad social, la confianza compartida, la reciprocidad en las acciones de otros agentes. Correspondientemente, experimentarían con desagrado las divergencias y el alejamiento de estas prácticas. Los motivos morales surgen en la conciencia de los agentes humanos de manera prácticamente imperceptiblemente para ellos, a pesar de su carácter convencional y artificial. No obstante, la fuerza de tales motivos se experimenta como *necesaria*, como si fuera una relación natural entre el agente y la obligación, pero tenemos que concluir que no es sino una ficción de nuestra psicología:

Por tanto, nos imaginamos que existe un nuevo acto mental, al que denominamos volición de una obligación, y nos figuramos que la moralidad depende de ello. Pero ya hemos probado que no existe tal acto y que, por consiguiente, las promesas no imponen ninguna obligación natural (T 3.2.5.12 SB 523).

Con «acto de la mente», en este contexto, Hume parece referirse a un acto *natural*, no al sentimiento moral que un agente *de hecho* experimenta con la volición de hacer una promesa y que, al mismo tiempo, es el motivo que tiene para considerarla como una obligación que recae sobre él mediante el acto mismo de tener una volición. Desde el punto de vista de que las promesas son un «mecanismo normativo artificial, inventado para la convivencia y provecho de la sociedad», Hume afirma además que su “fuerza no difiere esencialmente de cualquier otro motivo de esperanza o miedo que pueda inducirnos a comprometernos con nuestra palabra y hacer que nos sometamos a alguna obligación” (T 3.2.6.15 SB 525). Así, las constricciones morales, creadas artificialmente mediante convenciones, se ejercen como coerciones externas al agente, de modo similar a cómo un factor externo, como una amenaza o un cierto anhelo, ejercen coerción sobre los agentes y los influyen a actuar, si bien de manera diferente. No es lo mismo actuar por miedo, sometido a una amenaza, que por un anhelo, como una recompensa que esperamos recibir en el futuro. La experiencia fenomenológica de estas formas de coerción es distinta, si bien su influencia sobre la voluntad de los agentes es similar. Respecto de la coerción moral, según las ideas generales de Hume, las notas fenomenológicas que la caracterizan son el placer o el dolor que constituyen al sentimiento moral. A continuación clarificaré qué papel tienen estas entidades en la teoría moral de Hume.

6.2. La utilidad, la amabilidad y la belleza en el naturalismo de Hume

Según el desarrollo de las ideas de Hume hasta aquí, parece que la utilidad, la amabilidad o la belleza no son entidades naturales, esto es, ni son ellas mismas objetos o propiedades de otros objetos naturales. Sin embargo, explicar la relación que de *hecho* guardan con objetos del mundo natural plantea retos teóricos importantes. Consideremos para empezar el siguiente pasaje, donde Hume nos dice que la belleza no es ninguna propiedad del mundo natural:

Euclides ha explicado plenamente todas las cualidades del círculo; pero en ninguna proposición ha dicho una palabra sobre su belleza. La razón es evidente. La belleza no es una cualidad del círculo. No está en ninguna parte de la línea cuyas partes se encuentren equidistantes de un centro común. Es sólo el efecto que esa figura produce en la [292] mente, cuya constitución particular la hace susceptible de experimentar tales sentimientos. En vano buscaríais la belleza en el círculo, o la rastrearíais mediante vuestros sentidos o mediante razonamientos matemáticos en todas las propiedades de esa figura (IPM Ap. I SB 291-292)

La belleza, de acuerdo con esta observación de Hume, es una cualidad de la mente derivada de la impresión de ciertas propiedades del círculo. Esto es, no es reducible a alguna o algunas de las propiedades del círculo y tampoco es la suma de todas sus propiedades. Ahora bien, las ideas

anteriores que hemos examinado de Hume permiten, me parece, afirmar que la percepción de la belleza, en los términos de esta cita, es del tipo de los colores o los sabores, esto es, consiste en una propiedad secundaria. Desde el punto de vista de un agente individual, esta perspectiva requiere que de alguna manera sea educada o entrenada (como una habilidad), de manera que su disposición a percibir tal tipo de propiedad sea potenciada y afinada. Del mismo modo que sucede con la apreciación de la música, los sabores, los matices de colores, etc. En principio, todos los agentes humanos tienen disposiciones naturales para este tipo de impresiones, pero el placer o desagrado que obtienen con ejemplares diversos de ellas puede ser orientado y dirigido. Muchas instituciones sociales intervienen en este proceso, como el papel de familiares, autoridades especializadas, ejemplos socialmente reconocidos, etc. Creo que esta interpretación cobra fuerza si consideramos la analogía que establece Hume entre la belleza moral y la belleza natural, justo en el párrafo anterior al citado arriba:

Esta doctrina se hará todavía más evidente si comparamos la belleza moral con la natural, con la que guarda un estrecho parecido en muchos detalles. Toda belleza natural depende de la proporción, relación y posición de las partes; pero resultaría absurdo inferir de ello que la percepción de la belleza, como la de la verdad en los problemas de la geometría, consiste enteramente en la percepción de relaciones y se realiza enteramente mediante el entendimiento o las facultades intelectuales. En todas las ciencias nuestra mente investiga las relaciones desconocidas partiendo de las conocidas. Pero en todas las decisiones del gusto o belleza externa, todas las relaciones se hacen patentes a la vista de antemano; y a partir de entonces procedemos a experimentar un sentimiento de complacencia o de disgusto de acuerdo con la naturaleza del objeto y la disposición de nuestros órganos (IPM Ap. I SB 291-292).

Según esto, la percepción de la belleza, moral o natural, depende de algunas propiedades de los objetos, como la proporción de las partes del círculo entre sí, y de la disposición de nuestros órganos, pero éstos tienen que estar dirigidos o *educados* previamente por las investigaciones de “las ciencias”. En otras palabras, se requiere que el trabajo colectivo, histórico y social producido por las ciencias sea de alguna manera internalizado por los agentes, los cuales pueden entonces percibir lo inmediato a los sentidos y algo más, como la belleza. Según las ideas de Hume, no hay nada misterioso en este proceso de internalización de estándares de belleza o morales. Para explicarlo, sólo necesita recurrir a la institución de la educación, si bien ésta es compleja. Recordemos cómo Hume recurre a la figura «un hombre de buen sentido y juicio», esto es, un sabio o alguna autoridad en cierta materia. Consideraciones similares pueden hacerse sobre la benevolencia y la utilidad. Vimos que este recurso ajusta bien con las ideas generales de la interpretación del observador ideal sobre los juicios morales (v. sección 4.3).

El observador ideal tiene un punto de vista imparcial y toda la información relevante para juzgar correctamente sobre un asunto dado. Para él los juicios del entendimiento, esto es, aquellos en los que confiamos para tener representaciones *correctas* de ciertos objetos o relaciones entre objetos, son susceptibles de ser corregidos o validados por diversos medios. Con todo, en la medida en que es «ideal», podemos suponer confiadamente que sus recursos de verificación y corrección son ilimitados y que, por lo menos, coinciden con la experiencia humana acumulada en un momento y situación dados. Esto es importante porque tal observador no tiene que coincidir con ninguna personificación particular. Puede ser perfectamente impersonal, porque lo que esperamos de él o ella es que favorezca que los agentes adopten una actitud hacia sus juicios o sentimientos. El observador ideal predispone a los agentes a examinar mediante criterios de corrección sus estándares de actuación y pensamiento, a ajustar su experiencia con normas de apreciación y sus valoraciones con criterios reconocidos como valiosos por dicho punto de vista.

Quizá la situación más común en la que son requeridos estos procedimientos de corrección es el caso de conflicto entre juicios, por ejemplo, cuando dos agentes difieren en su apreciación de cierto juicio: uno los considera correctos y otro incorrecto. En estos casos, agentes con una actitud razonable hacia el conflicto pueden albergar dudas sobre la validez de sus juicios y entonces apelar a alguno de dichos procedimientos. En palabras de Hume, estos agentes podrían proceder así:

Quando surgen estas dudas no puede hacerse más que lo que se hace con otras cuestiones discutibles que se someten al entendimiento. Hay que utilizar los mejores argumentos que se sea capaz de idear; hay que reconocer que en algún lugar existe una norma verdadera y decisiva para el ingenio, la existencia real y el sentido común, y hay que tener indulgencia con quienes difieran en cuanto a la aplicación de esta norma (DNG §25, 231-232).

Esto es consistente con la idea de Wiggins que hemos considerado antes (v. arriba sección 4.3): la concepción de Hume de los juicios morales, aun considerando que son sentimientos, da lugar a la influencia de una actitud reflexiva y razonable hacia los casos en disputa. Según Hume, en efecto, un agente en una situación así «simplemente hace lo mismo que en el caso de los conflictos similares que son objeto del entendimiento». Así, el juicio moral parece tener las propiedades que lo hacen susceptibles de ser tratado como si fuera un juicio del entendimiento o una representación de objetos o relaciones entre ellos. Hay, sin embargo, una diferencia fundamental: los juicios morales no pueden ser verdaderos, a diferencia de los producidos por el entendimiento. Quiero terminar esta sección examinando si esta última afirmación hace de la teoría de los juicios morales que estamos examinando un tipo de «teoría del error».

De acuerdo con esta teoría, propuesta por Mackie en su *Ethics: Inventing Right and Wrong* (Mackie 1997/2000),¹⁰⁴ los juicios morales son sistemáticamente falsos porque parecen referirse a entidades o cualidades o relaciones de muy extraña naturaleza, enteramente distintas a cualquier otra realidad del universo. Ésta es la tesis de la «rareza» [*querness*] de las supuestas entidades morales. Además, Mackie considera que si tales entidades existieran, para que los juicios morales fueran verdaderos necesitaríamos de alguna facultad mental peculiar, distinta a las facultades de los sentidos, para percibirlos (cf. Mackie 2000, p. 42). Como parece imposible que estas dos condiciones puedan ser satisfechas, entonces los juicios morales son sistemáticamente falsos. A pesar de esto, sí poseen fuerza constrictiva sobre los agentes, la cual reside en su valor práctico. Por ejemplo, permite a agentes egoístas alcanzar sus fines en condiciones en las que tienen que tomar en cuenta los fines e intereses de otros agentes. Para Mackie éste no es sino «el valor de un lazo o ligadura invisible y en realidad ficticio», es decir, en el fondo es sólo una disposición psicológica adoptada por los agentes pero que carece realmente de toda realidad fuera de que los agentes la aceptan –es tan absurda como la idea de obligación natural que rechaza Hume (ref.). Según esta idea, la fuerza de los juicios morales no es más real que el caso del honor militar o la lealtad a los camaradas o a una organización (cf. Mackie 2000, p. 326 y ss). Esta explicación es consistente con la interpretación *sentimentalista* que hace Mackie de la teoría de los juicios morales de Hume. Según la cual, para Hume los juicios morales son *meros* sentimientos sin valor cognitivo alguno. Estoy de acuerdo con Mackie en que los juicios morales humeanos no pueden ser verdaderos y que se fundan en sentimientos, pero creo que una interpretación más interesante y plausible de los mismos, como he defendido, la ofrece la noción del observador ideal.

Hasta ahora he aportado elementos para defender la interpretación de una teoría del observador ideal (v. §4.3). Según esta teoría, los juicios morales de Hume son en efecto sentimientos valorativos que se refieren a una entidad como el carácter de uno mismo o de otra persona considerado de manera general o imparcial. De modo que no requieren referirse a ninguna propiedad del mundo natural. Ahora, cuando hablamos del valor cognitivo de los juicios morales creo que el enfoque de Mackie –y de Hume– es el correcto: queremos saber de qué entidades hablan tales juicios y qué conocimiento nos aportan de ellas. Con todo, Mackie parece pedir que tales entidades se ubiquen en el mundo natural. Sólo así podríamos aceptar que los juicios morales

¹⁰⁴ Mis referencias son a la edición en español de 2000.

tienen valor cognitivo. Ante esto, Hume podría no ser este tipo de naturalista. Según hemos visto, para él los juicios morales se refieren a ciertas entidades, como el carácter abstracto de una persona, pero también a aspectos más particulares del mismo, como su benevolencia, amabilidad, humanidad, desinterés, celo, etc., y también la utilidad de un carácter o de una acción, la belleza de las acciones virtuosas o útiles, etc. Según él, la percepción de estas entidades la obtenemos por medio de la imaginación, el hábito y la costumbre. Esta facultad y las demás disposiciones tienen valor cognitivo en la medida en que informan a los agentes y orientan su valoración de las mismas. Además, en mi opinión, no tiene nada de misteriosas, por contraste con la *intuición* que, según Mackie, se requeriría para percibir a las entidades morales, si es que hay alguna (cf. Mackie 2000, 42 y ss). Las ideas de Hume sobre la imaginación, por ejemplo, son que ésta no nos da conocimiento representacional del mundo, pero sí nos permite elaborar en alguna medida tal tipo de conocimiento a partir de ciertas impresiones sensibles, por ejemplo, el conocimiento de las relaciones causales. Según hemos visto, este conocimiento es de gran importancia práctica, si bien difícilmente podríamos aceptarlo como conocimiento en toda regla. En otras palabras, aun cuando positivamente no podamos aceptar las creencias derivadas por la imaginación, su valor práctico nos es imprescindible para prácticamente todo lo que hacemos (ref.). Si la teoría de Mackie es correcta sobre las entidades *raras*, entonces incluso las creencias causales tendrían que ser todas ellas sistemáticamente falsas, y es parece que esto no es el caso. En efecto, parece que una teoría del juicio verdadero tan fuerte como la que presupone Mackie no podría admitir a las creencias causales que la teoría de Hume ofrece. Además, se puede elaborar una respuesta a esta posición de Mackie con base en la idea de Hume en DNG acerca de que los juicios morales son más constantes o estables que los juicios del entendimiento.

La idea, que cité antes, es la siguiente “No son lo mismo los principios morales que las opiniones especulativas de cualquier tipo. Éstas (*sic*) últimas están en cambio y revolución continuos. El hijo se adhiere a un sistema distinto que el padre. Es más, apenas hay alguien que pueda presumir de gran constancia y uniformidad a este respecto” (DNG §33, 235; v. arriba p. 141). Según esto, hay más inconstancia y duda en el conocimiento especulativo o teórico que en el práctico. Si esto es correcto, parece que podríamos decir con base en ello que las entidades del mundo natural, estudiadas por dicho conocimiento, son más “raras” que las entidades del mundo moral. De éstas, por contraste, tenemos conocimiento certero y no suscitan los conflictos que aquellas. Con todo, me parece que esta opinión de Hume no es plausible —en todo caso, parece

haber tantos conflictos y diferencias en el conocimiento relativo a cualquiera de los dos mundos mencionados. A pesar de esto, creo que es una buena opinión para mostrar que también la designación de «entidades raras» de Mackie es impugnable: a él le parecen raras las entidades morales, pero casi cualquier que lea o vea alguna película sobre los animales abismales, el mundo de los microorganismos, o la inteligencia de los delfines o incluso las costumbres de los gatos domésticos puede experimentar muy vivamente que son entidades muy raras. El mundo natural puede ser, para una mente curiosa, muy raro.

Ahora bien, si los procesos de la imaginación son de utilidad práctica para la formulación de creencias factuales, entonces podemos suponer que también lo son para las creencias y los juicios morales. De acuerdo con esta analogía, la utilidad moral, la virtud y la belleza morales no serían más o menos raras que la causalidad. El proceso de derivación de la creencia causal a partir de la percepción sensible de ciertas entidades y sus propiedades no sería tampoco más raro que el proceso de derivación de las entidades morales mencionadas a partir de la percepción sensible de otras entidades y sus propiedades. Por otra parte, la interpretación del observador ideal encuentra otros elementos a su favor en la noción de prácticas morales para explicar a los juicios morales.

6.3. La teoría del observador ideal de Hume y las prácticas morales

Según he dicho, la teoría del observador ideal consiste en la afirmación de que las distinciones morales están basadas en «los sentimientos hipotéticos de un espectador ideal (un agente al que suponemos plenamente informado, enteramente objetivo y no auto-interesado)». Tanto en el capítulo cuarto como la sección precedente, hemos estudiado cómo este observador ideal es internalizado por cada agente individual por medio de un cierto proceso de educación, de entrenamiento de la imaginación y de adoctrinamiento de valores y normas de apreciación. Este proceso explica las funciones del hábito y la costumbre, las actitudes que contribuyen a conformar a la imaginación a seguir pautas específicas de formación y derivación de creencias. Quizá la característica más importante de este proceso es que a través de él cada agente individual se convierte en miembro funcional en una comunidad de agentes que dan respuestas consistentes a estímulos similares. Los agentes humanos requieren de una comunidad para percibir, evaluar, conocer, imaginar, etc. Son parte de un colectivo que no se reduce a la suma de sus individuos, pero tampoco cada individuo expresa totalmente al colectivo ni a sus condiciones epistemológicas

y normativas, entre las que me he concentrado en las morales. Con base en esto, me parece, pueden establecerse relaciones interesantes con la noción de *práctica moral*.

Siguiendo ideas de Hume, voy a caracterizar a una práctica moral como al conjunto de pensamientos, acciones, actitudes, *dispositivos* de alabanza y censura dirigidos por algunas reglas y principios morales por parte de grupos de agentes relativamente similares entre sí –al menos respecto de tales conjuntos. Ni las reglas ni los principios tienen que estar formulados explícitamente. Pueden ser sólo ciertas regularidades admitidas tácitamente por los usuarios de la práctica o ciertas ideas muy generales que no suelen ponerse en duda. Ésta es una caracterización muy amplia, pero creo que tiene ventajas mantenerla así para dar lugar dentro de ella a varios elementos de la compleja vida social. Me parece, además, que partes importantes de la vida social son coextensivas con conjuntos específicos de las prácticas morales. No hay necesidad de suponer que todas las sociedades humanas son homogéneas ni tampoco de delimitar claramente el grado en que hay diferencias entre ellas. Lo que, por ahora, me interesa es destacar que si esta caracterización de las prácticas morales es plausible, entonces dicha noción parece implicar que socialmente puede concebirse que grupos diferentes de agentes puedan coincidir en sus prácticas morales o que podría haber traslapes interesantes entre diferentes partes de sus prácticas morales. Por ejemplo, supongamos dos grupos de agentes que se distinguen entre sí por su actitud hacia el honor. Para uno de estos grupos ésta es una virtud preponderante y cada agente percibe el daño a su honor como un daño personal importante y fundamental para la idea de sí mismo. Para el otro grupo de agentes el honor no es tan importante como la fama. Para éste, la fama tiene preponderancia de manera que incluso si dañar al honor propio contribuye a la fama personal, esto no es percibido como un daño personal, sino en todo caso al contrario. Por lo demás, ambos grupos coinciden en la regla de ayudar a los necesitados. Sólo que esta regla se subordina al honor para los primeros y a la fama para los segundos y aun así ayudar a los necesitados es una regla moral que ambos grupos tienen en alta estima. Este ejemplo esquematiza cómo en una misma sociedad podría haber grupos de personas con prácticas morales diferentes entre sí respecto de ciertos aspectos morales y, al mismo tiempo, con otros aspectos morales similares entre ellos. Antes de hablar de cómo se relaciona esta noción de las prácticas morales con la teoría del observador ideal necesito agregar otro elemento derivado de las ideas de Hume.

Hasta donde puedo apreciar que las ideas de Hume son compatibles con esta noción de las prácticas morales, creo que su atención se dirigiría a las *semejanzas* que hay entre las prácticas de

distintos grupos sociales. Su metodología genealógica lo conduciría a establecer en la naturaleza humana una base sólida para explicar las semejanzas entre prácticas humanas y, quizá, sobre todo para sostener que las diferencias apreciables no son suficientes para descartar que, en el fondo, se trata siempre de la manifestación de la moral humana. Para Hume, la moral humana es coextensiva con la sociedad misma, de modo que no puede entenderse una sin la otra. Como hemos visto antes, las sociedades humanas han surgido porque ciertos acuerdos morales básicos se han celebrado y cuya finalidad es fortalecer a los elementos morales característicos de la naturaleza humana. Para Hume, la sociedad es una extensión de la naturaleza moral humana. En consecuencia, las diferencias entre prácticas morales no son importantes, excepto como expresión de la complejidad misma de las sociedades humanas. Las diversas condiciones que han intervenido para formar a las sociedades, relativas a la geografía, la tradición, la historia, los retos de sobrevivencia, etc., explican esta complejidad y las diferencias entre las prácticas morales. Pero no podemos tomar a esta complejidad para negar la moralidad básica necesaria para el funcionamiento de las sociedades humanas. Por otra parte, aunque las prácticas morales están, de alguna manera, en el *origen* de las sociedades humanas, su fuerza cohesionadora ejerce sin cesar su influencia.

Las reglas y principios morales que caracterizan a las prácticas morales mantienen su carácter rector de la conducta de sus usuarios de manera constante. Podemos reconocer sus efectos en cada ocasión en que nuestra conducta sigue las pautas sociales aceptadas. Cuando sucede alguna divergencia, ésta, muchas veces, es sólo una reacción a tales pautas. En este sentido, son también un reconocimiento de su influencia. Por ejemplo, una comunidad de médicos comienza a mostrar dudas sobre el alcance del principio hipocrático de: “Venerar como a mi padre a quien me enseñó este arte, compartir con él mis bienes y asistirles en sus necesidades; considerar a sus hijos como hermanos míos, enseñarles este arte gratuitamente si quieren aprenderlo...” Las dudas podrían referirse a si es necesario en todos los casos cubrir los abusos o la negligencia cometidos por colegas médicos. Tal vez se ha llegado a un punto en que para muchos médicos de una cierta comunidad la idea de que discutir abierta y ampliamente los fallos de la profesión trae más beneficios para su comunidad que mantenerlos en secreto. Esto puede ser visto simplemente como un reconocimiento de que las relaciones de la práctica médica con otras prácticas sociales, como las prácticas jurídicas o las complejas políticas administrativas de los hospitales actuales, requieren de tal apertura. O pueden ser la expresión del reconocimiento de que la atención y el respeto para los pacientes es de más peso que la secrecía profesional. Cualquiera que sea el caso

(la estima de cómo se obtiene el mayor beneficio para la comunidad moral o una percepción novedosa de las complejidades de la práctica dada por sus relaciones con otras prácticas o un cambio en la percepción de cuál es el principio moral preponderante), estas divergencias con el estado de la práctica en un momento anterior pueden interpretarse como una afirmación de lo que realmente importa para la comunidad: su beneficio global, su compatibilidad con otras prácticas o su respeto a principios morales fuertes. Como decía antes, las diferencias son menos importantes que el reconocimiento para los usuarios de una práctica de que hay un interés moral fundamental en ajustar su conducta a pautas compartidas por todos. Ésta es la expresión de la moral y captura un elemento importante relativo a la universalidad de las regulaciones morales. *Todos* los agentes integrantes de una cierta práctica, en efecto, ajustan su conducta al estándar moral compartido, al menos en la medida en que entienden su propia participación en la misma. Creo que este punto permite ahora vincular estas ideas con la teoría del observador ideal.

Creo que podemos entender que la relación de los agentes individuales con las reglas y principios de la comunidad moral a la que pertenecen se desarrolla en dos niveles: el nivel de la práctica normal o habitual y el nivel de la reacción o ajustes a las condiciones normales de la práctica. En el primer nivel los estándares regulativos de la práctica podrían ser incluso imperceptibles para los agentes, excepto por los peculiares sentimientos placenteros que hemos considerado antes, como la utilidad y la belleza. Hume se refiere explícitamente a esta idea discutiendo los motivos que impulsan a los agentes a actuar de acuerdo con los deberes morales. Creo que la idea, en general, puede aplicarse a cualquier contenido moral regulativo:

Y sin embargo, ¿acaso no es posible que el sentimiento de moralidad, o de deber, produzca una acción sin necesidad de otro motivo? Respondo: es posible, pero ello no supone objeción alguna a la presente doctrina. Cuando un motivo o principio de virtud es común a la naturaleza humana, la persona que siente faltar en su corazón ese motivo puede odiarse a sí misma por ello y realizar la acción sin la existencia del motivo, basándose en un cierto sentido del deber y con la intención de adquirir con la práctica ese principio virtuoso, o al menos para ocultarse a sí misma en lo posible la ausencia de dicho motivo. Al hombre que no siente realmente gratitud en su interior le agrada, sin embargo, realizar acciones de agradecimiento, pensando que de esa forma ha cumplido con su deber (T 3.2.1.8 SB 479-480).

Hume dice aquí que quien no cumple con sus deberes morales «puede odiarse a sí mismo», esto es, experimentar un vivo desagrado hacia su propia persona, o un sentimiento de desaprobación de sí misma (cf. Rivera, 86). Como en este pasaje en general examina que los agentes pueden actuar sólo por alguna idea del deber, sin que necesariamente experimenten los sentimientos placenteros de belleza o de utilidad, parece seguirse que en general las acciones reguladas

moralmente son marcadas fenomenológicamente por el placer mencionado, relativo a la belleza o utilidad de sus acciones. Incluso, las acciones realizadas por deber son, además, impulsadas por esa belleza y si el agente no la experimenta entonces puede apropiarse de ella “con la intención de adquirir con la práctica ese principio virtuoso” y que eventualmente logre conformar sus actos con los de ciertas prácticas (cf. Wiggins 2006, 78). En otras palabras, hay un importante elemento motivador para que un agente dado alcance los estándares regulativos de una cierta práctica mediante la internalización de los sentimientos placenteros, favorables, de las conductas reguladas y sancionadas por dichas prácticas. La belleza de las acciones virtuosas o reguladas moralmente es, para Hume, el más fuerte de estos motivos. Según presenté antes con la explicación sobre los juicios morales (v. §4.3), este proceso de internalización pasa por la idea de un observador ideal.

El tratamiento del segundo nivel de la relación entre un agente individual y las reglas propias de una cierta práctica muestra otros aspectos de cómo opera el observador ideal en la formación de los motivos virtuosos. Este nivel se refiere a esas condiciones en que los agentes reaccionan o ajustan su conducta a las condiciones normales de la práctica, el primer nivel. El punto de Hume, destacado por Wiggins, es que los agentes se apropian del sentimiento de belleza para adquirir los motivos por los que regulan su conducta según los estándares morales. Pero esto requiere, me parece, que los agentes reconozcan su interés por los estándares morales en cuestión como propios o característicos de su propia idea de persona. Por ejemplo, que reconozcan que su posición como médicos, en tanto que médicos, requiere de ajustar su conducta a los estándares que contravienen la idea de la secrecía profesional –según el ejemplo presentado arriba. La idea de la persona que se desea ser, en tanto que médico, funge como criterio normativo para ajustar la conducta a ella y es, me parece a mí, una forma del observador ideal. En este caso, el observador ideal no es una herramienta de la deliberación práctica sino, en términos más generales, de los motivos morales que marcan la conducta apropiada de una persona. Desde el punto de vista de Hume, la sola idea o creencia no son suficientes para motivar. Se requiere del sentimiento correspondiente. Es decir que la mera idea de la persona que se desea ser no es suficiente para motivar ninguna conducta al respecto. La motivación al respecto se activa cuando se percibe tal idea bajo un sentimiento placentero, esto es, bajo la sensación de utilidad o belleza. Todo esto nos da una idea muy precisa de la naturaleza de la motivación moral para Hume.

Para Hume, las reglas morales no son meramente *mandatos*, sino, formas de experimentar y evaluar la realidad social, desde un punto personal y en perspectiva social. Son así formas de

evaluación y conocimiento de acuerdo con una perspectiva social que es experimentada mediante sentimientos placenteros. Como lo ha puesto Wiggins, las reglas morales son «el modo según el cual los agentes evalúan las cosas con las que se encuentran y la norma que pueden seguir para que su conducta sea responsable [*answerable*]. La regla abstracta encarna o condensa perspectivas e intereses generales y establecidos en común. Pero el estándar o la regla misma es la creación de sentimientos y modos de pensar en los que todos participamos» (cf. Wiggins 2006, 48). En otras palabras, las reglas morales constituyen un tipo de experiencia de las realidades humanas y el tipo peculiar de reacciones de una comunidad a dichas realidades interiorizadas por cada agente individual y resultan ser motivacionalmente eficaces para regular conductas en la medida en que son experimentadas placenteramente.

Esta actitud se alcanza simplemente con la participación de los agentes en las diversas prácticas sociales. Los medios de verificación mencionados son interiorizados en cada agente por el complejo proceso educativo que nos permite compartir los diversos estándares de percepción, apreciación y corrección. La teoría del observador ideal le da forma a este proceso de manera muy plástica, pero en realidad, según esta interpretación, tal punto de vista social opera en nosotros como una especie de plataforma psicológica que sostiene la vida mental de cada individuo agremiado a la sociedad. Así, en la medida en que las prácticas morales son coextensivas con las diversas prácticas sociales y, quizá, con la sociedad misma, podemos apreciar con más claridad la idea de Wiggins, citada arriba, respecto de cómo la virtud moral, esto es, la conducta humana que expresa las reglas morales bajo la forma de un sentimiento de aprobación personal, nos aparece, desde el punto de vista de Hume, «como el más fino florecimiento posible de las naturalezas originales y adquiridas de los agentes humanos» (cf. arriba p. 186).

7. Bibliografía

- Alanen, L. (2005), "Reflection and Ideas in Hume's Account of the Passions" en Baier, A., Jenkins, J., Whiting, J., & Williams, C. (Eds.). (2005). *Persons and Passions: Essays in Honor of Annette Baier*. Notre Dame: University of Notre Dame Press.
- Alanen, L. (2006), "The Powers and Mechanisms of the Passions" en Traiger, Saul (ed.) (2006), *The Blackwell Guide to Hume's Treatise*, Blackwell, Oxford.
- Alston, W. P. (1968), "Moral attitudes and moral judgments", *Nous*, 1-23.
- Anscombe, G. E. M. (1986), 'Del razonamiento práctico' (Tr. J. J. Utrilla) en J. Raz, (ed.), *Razonamiento práctico*. México: FCE.
- Árdal, P. S. (1989), "Hume and Davidson on Pride", *Hume Studies*. Vol. XV, n. 2, noviembre.
- Audi, R. (2006), *Practical Reasoning and Ethical Decision*. N.Y.: Routledge.
- Ayer, A. J. (1994), "Crítica de la ética y de la teología" en su *Lenguaje, verdad y lógica*. (Tr. M.Suárez). Barcelona: Planeta-Agostini (col. Obras maestras del pensamiento contemporáneo).
- Baier, A. (1990), "What Emotions are About?", *Philosophical Perspectives*, Vol. 4, Action Theory and Philosophy of Mind (1990), pp. 1 -29
- Baier, A. (1991), *A Progress of Sentiments. Reflections on Hume's Treatise*. Cambridge: HUP.
- Baier, A. (1993), "Moralism and Cruelty: Reflections on Hume and Kant", *Ethics*, 103, 436-57.
- Baillie, J. (2000), *Hume on morality*. London y New York: Routledge.
- Bennett, J. (2001), *Learning from Six Philosophers Vol. 2*. Oxford: Clarendon Press.
- Biro, J. (1993), "Hume's new science of the mind" en Fate Norton, D, (ed.), *The Cambridge Companion to Hume*. Cambridge, CUP.
- Botros, S. (2006), *Hume, Reason and Morality. A Legacy of Contradiction*. Londres y NY: Routledge.
- Brand, W. (1992), *Hume's theory of moral judgment: a study in the unity of A Treatise of Human Nature*. Kluwer Academic Publishers: Dordrecht.
- Broughton, J. (2006), "Impressions and Ideas" en Traiger, Saul (ed.) 2006: *The Blackwell Guide to Hume's Treatise*.
- Byrne, R. (2007), *The Rational Imagination. How people create alternatives to reality*. Cambridge, Mass.: MIT Press.
- Calhoun, Ch. y Solomon, R. (eds.) (1989), *¿Qué es una emoción? Lecturas clásicas de psicología filosófica*. México: FCE.

- Copp, D. (2005), *The Oxford handbook of ethical theory*. Oxford University Press, New York.
- Cullity, G. and Gaut, B. N., (1997) *Ethics and practical reason*. Oxford N.Y.: Clarendon Press y OUP.
- Currie, G., (2002) “Imagination as Motivation”, *Proceedings of the Aristotelian Society*. New Series, vol. 102 pp. 201-216.
- D’Arms, J. y Jacobson D. (2003), “The significance of recalcitrant emotion (or, anti-quasijudgmentalism”, en Hatzimoysis, A. (ed.) (2003) *Philosophy and the Emotions*. CUP.
- Damasio, A. (2008), *El error de Descartes. La emoción, la razón y el cerebro humano*. (Tr. J. Ros). Barcelona: Crítica (1ª ed. Colec. Drakontos Bolsillo).
- Davidson, D. (1976), “Hume’s cognitive theory of Pride”, *The Journal of Philosophy*, vol. 73, no. 19. Nov. 4. (Traducción al español en: Davidson, D., (1995) en *Ensayos sobre acciones y sucesos*. Barcelona, UNAM-Crítica.)
- Davidson, D., (1995), “Acciones, sucesos y causas”, en *Ensayos sobre acciones y sucesos*. Barcelona, UNAM-Crítica.
- De Sousa, R. (2013), “Emotion”, *Stanford Encyclopedia of Philosophy*. (Spring 2013 Edition), Edward N. Zalta (ed.), forthcoming URL = <<http://plato.stanford.edu/archives/spr2013/entries/emotion/>>
- Deigh, J. (1994), “Cognitivism in the Theory of Emotions”, *Ethics*, 824-854.
- Fate Norton, D. ed. (1993), *The Cambridge Companion to Hume*. Cambridge: CUP.
- Fodor, J. (2003), *Hume variations*. Oxford: OUP.
- Garret, D. (2006), “Hume’s conclusion” en Traiger, Saul (ed.) (2006), *The Blackwell Guide to Hume’s Treatise*.
- Gaskin, J. C. A. (1993), “Hume on religion”, en Fate Norton (ed), (1993), *The Cambridge Companion to Hume*.
- Gazzaniga, M. S. (2008), *Human. The Science Behind What Makes Us Unique*. N.Y.: Harper Collins.
- Gibbard, A. (1990), *Wise Choices, Apt Feelings*. Cambridge, Mass.: CUP.
- Gill, M. B. (2000), “Hume’s Progressive View of Human Nature”, *Hume Studies*, abril (pp. 87-108).
- Guyer, P. (ed.) (1998), *Kant’s Groundwork of the Metaphysics of Morals: Critical Essays*. Lanham, MD: Rowman and Littlefield
- Hampton, J. (1995), “Does Hume Have an Instrumental Conception of Practical Reason?” *Hume Studies*, Vol. XXI, N. 1, april. Pp. 57-74.
- Hansberg, O. (1996), *La diversidad de las emociones*. México: FCE.
- Hardin, R. 2007, *David Hume: Moral and Political Theorist*. Oxford: OUP.

- Hausman y Hausman (1997), *Descartes Legacy: Minds and Meaning in early modern philosophy*. Toronto: University of Toronto Press.
- Hume, D. “De la norma del gusto” en Hume (2011), pp. 219-237.
- Hume, D. (1978, 2a ed.), *A Treatise of Human Nature*, ed. por Selby-Bigge. Oxford: Clarendon Press.
- Hume, D. (1988), *An Enquiry Concerning the Principles of Morals*, ed. por T. L. Beauchamp, Oxford: OUP.
- Hume, D. (1988), *Tratado de la naturaleza humana* (Tr. y ed. de F. Duque). Madrid: ed. Tecnos.
- Hume, D. (1999), *An Enquiry Concerning Human Understanding*, ed. por T. L. Beauchamp. Oxford: OUP. (Col. Oxford Philosophical Texts).
- Hume, D. (2000), *A Treatise of Human Nature*, edited by David Fate Norton and Mary J. Norton, Oxford/New York: OUP.
- Hume, D. (2008), *Ensayos morales y literarios* (Ed. y tr. de E. Trincado). Madrid: ed. Tecnos.
- Hume, D.(2011), *Ensayos morales, políticos y literarios* (Ed. de E. Miller. Tr. de C. M. Ramírez. Madrid: Ed. Trotta.
- Johnson, M. (1993), *Moral imagination. Implications of Cognitive Science for Ethics*. Chicago: The University of Chicago Press.
- Jones, P. (1993), “Hume’s literary and aesthetic theory” en Fate Norton (ed.), (1993).
- Justin D’Arms y Jacobson D. (2005), “Sensibility Theory and Projectivism”, en Copp, D. (2005), *The Oxford Handbook of Ethical Theory*.
- Kant, I. (1797), *The Metaphysics of Morals* en Kant, *Practical Philosophy* (Tr. y ed. M. J. Gregor). Cambridge: CUP.
- Kant, I. (1948), *Groundwork of the Metaphysic of Morals*, ed. y tr. H. J. Paton, London: Hutchinson & Co. 1948.
- Kant, I. (2001), *Crítica de la razón práctica*, ed. y tr. D. M. Granja Castro, México: FCE y Miguel Angel Porrúa.
- Karlsson, M. M. (2006), “Reason, Passion, and the Influencing Motives of the Will”, en Traiger, Saul (ed.) (2006), *The Blackwell Guide to Hume’s Treatise*.
- Kenny, A. (1963), “Action” en *Emotion and Will*. London: Routledge.
- Korsgaard, C. M. (1996), "Skepticism about practical reason" en *Creating the Kingdom of Ends*. Cambridge, New York: CUP.
- Korsgaard, C. M. (1999), “The General Point of View: Love and Moral Approval in Hume’s Ethics”, *Hume’s Studies*, abril-noviembre XXV: 1-2.
- Lenman, J. (2014), “Moral Naturalism”, *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2014/entries/naturalism-moral/>>.

- Levi, W. A. (1964), "The Two imaginations", *Philosophy and Phenomenal Research*, vol. 25, no. 2, Dec. pp. 188-200).
- Lewis, D. (1969), *Convention*. Cambridge: HUP.
- Lovibond, S. (1983), *Realism and Imagination in Ethics*. Oxford: Basil Blackwell pub.
- Lowe, E. J. (2004), *An Introduction to the Philosophy of the Mind*. CUP (Ebook).
- Mackie, J. (1977), *Ethics: Inventing Right and Wrong*. London: Penguin. (En español: Mackie, J. L. (2000). *Ética: La invención de lo bueno y lo malo*. Gedisa).
- Mackie, J. (1980), *Hume's Moral Theory*. London: Routledge & Kegan Paul.
- MacLachlan, Ch. (1986), "Hume and the Standard of Taste", *Hume Studies*, Vol. XII, 1.
- McDowell, John (1998), "Might There Be External Reasons?" en John McDowell (ed.), *Mind, value, and reality*. Cambridge, Mass.: HUP, pp. 95-111.
- McIntyre, J. L. (2006), "Hume's "New And Extraordinary" Account of the Passions", en Traiger, Saul (ed.) (2006), *The Blackwell Guide to Hume's Treatise*, Blackwell.
- McNaughton, D. y. R., Piers (2004), 'Duty, Rationality, and Practical Reasons', en A. y. R. Mele, Piers, (ed.), *The Oxford Handbook of Rationality*, Oxford University Press, New York, pp. 110-131.
- Mele, A. R., Rawling, P. (ed), (2004), *The Oxford Handbook of Rationality*. N. Y.: OUP.
- Merril, K. R. (2008), *Historical Dictionary of Hume's Philosophy*. The Scarecrow Press: Lanham, Toronto y Plymouth.
- Miller, A. (2003), *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press.
- Millgram, E. (1995), "Was Hume a Humean?", *Hume Studies*, Volume XXI, Number 1, Abril, pp. 75-93.
- Morris, W. E. (2006) "Belief, probability, normativity", en Traiger, Saul (ed.), *The Blackwell Guide to Hume's Treatise*.
- Morton, A. (2006), 'Imagination and Misimagination', en Shaun Nichols (ed.) (2006), *The Architecture of the Imagination*.
- Nagel, T. (2004) *La posibilidad del altruismo*. México: FCE.
- Nichols, S. (2004), *Sentimental Rules. On the Natural Foundations of Moral Judgments*. Oxford: OUP.
- Nichols, S. (ed.) (2006), *The Architecture of the Imagination. New Essays on Pretence, Possibility, and Fiction*. Oxford: OUP.
- Nuyen, (1988), "The Role of Reason in Hume's Theory of Belief", *Hume Studies* Vol. 2 Nov. 1988 372-389.
- Pigden, Ch. R. ed. (2009) *Hume on Motivation and Virtue*. Basingstoke: Palgrave Macmillan.
- Platts, M. (1988), "Hume and Morality as a Matter of Fact", *Mind*, (Abr. 1988).

- Platts, M. (1999), *Sobre usos y abusos de la moral. Ética, SIDA, sociedad*. México: Ed. Paidós-IIF.
- Radcliffe, E. S. (1994), "Hume on Motivating Sentiments, the General Point of View, and the Inculcation of "Morality", *Hume Studies*, Vol. XX, Núm. 1, Abril. Pp. 37-58.
- Radcliffe, E. S. (ed). (2008) *A Companion to Hume*. Malden, Oxford, Carlton: Blackwell Publishing.
- Rawls, J. (2000), *Lectures on the History of Moral Philosophy*, ed. Barbara Herman. Cambridge, MA: HUP.
- Raz, J. (1986), (ed.), *Razonamiento práctico*, FCE, México.
- Rescorla, M. (2011), "Convention", *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2011/entries/convention/>>.
- Rivera, F. (2006), "Hume y la obligación moral" en Platts, M. (comp.) *Conceptos éticos fundamentales*. México: UNAM.
- Rolls, E. T. (2007), *Emotion Explained*. Oxford: OUP.
- Russell, L., (2009) "Two Kinds of Normativity: Korsgaard vs. Hume", en Pigden, Charles R. (ed.) (2009), *Hume on Motivation and Virtue*.
- Ryle, G. (2005), *El concepto de lo mental* (Tr. E. Rabosi). Barcelona: Ed. Paidós. (London: Hutchinson, 1949).
- Saab, S. (1999). "La creencia", en Luis Villoro (ed.), *El conocimiento*. Madrid: ed. Trotta. Vol. 20 de la Enciclopedia Iberoamericana de Filosofía.
- Scheffler, S. (1994), *Human Morality*. Oxford: OUP.
- Scheler, G. F. (1995) *Desire. Its Role in Practical Reason and the Explanation of Action*. Cambridge MA: MIT Press.
- Schroeder, T. and Carl Matheson (2006), "Imagination and Emotion", en Nichols S. (ed.) (2006), *The Architecture of the Imagination*.
- Setiya, K. (2007), *Reasons without Rationalism*. Princeton: PUP.
- Smith, M. (1987), "The Humean Theory of Motivation", *Mind*, Vol. 96, No. 381 Jan.
- Smith, M. (1994), *The Moral Problem*, Oxford: Blackwell.
- Smith, M. (2004), "Humean Rationality", en Mele, A. R., Rawling, P. (ed), (2004), *The Oxford Handbook of Rationality*.
- Strawson, G. (2003), *The Secret Connexion. Causation, Realism, and David Hume*. Oxford: Clarendon Press.
- Streminger, G., (1980), "Hume's theory of imagination", *Hume Studies*, vol. VI, number 2 Noviembre.
- Stroud, B., (1986), *Hume*. (Tr. A. Ziri6n). México: UNAM-IIF.

- Stroud, S. (2003), "Weakness of Will and Practical Judgement", en Stroud, Sarah y Tappolet Christine (2003), *Weakness of Will and Practical Irrationality* (eds). Oxford: OUP.
- Traiger, S. (ed.) (2006), *The Blackwell Guide to Hume's Treatise*. Oxford: Blackwell.
- Taylor, J. (2008), "Hume on Beauty and Virtue", en Radcliffe, E. S. (ed.), (2008).
- Velleman, J. D. (2000), *The Possibility of Practical Reason*. Oxford: Clarendon.
- Wallace, J. (1990), "How to argue about Practical Reason". *Mind* 99, 355-383.
- Wallace, J. (2001), "Normativity, Commitment, and Instrumental Reason", *Philosophers' Imprint*, Vol. 1, no. 3, dec.
- White, A. R. (1968), *The philosophy of action*. London: OUP.
- Wiggins, D. (2002), "A Sensible Subjectivism?", en *Needs, Values, Truth. Essays in the Philosophy of Law*. Oxford: Clarendon Press. 3a ed. corregida.
- Wiggins, D. (2006), "Hume's genealogy of morals", en *Twelve Lectures on the Philosophy of Morals*. Cambridge: HUP.
- Wilbanks, J. (1968), *Hume's theory of imagination*. The Hague: Martinus Nijhohh.
- Williams, B. (1993), "Razones internas y externas", en su *La fortuna moral. Ensayos filosóficos 1973-1980*. (Tr. S. Marín). México: UNAM-IIFs.
- Wright, G. H. v. (1983), *Practical reason*. Oxford: Basil Blackwell.
- Wright, J. P. (1983), *The Sceptical Realism of David Hume*. Minneapolis: University of Minnesota Press.