

UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO

PROGRAMA DE POSGRADO EN FILOSOFÍA
FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS
LÓGICA, FILOSOFÍA DEL LENGUAJE Y FILOSOFÍA DE LA MENTE

JUICIOS Y EMOCIONES MORALES: UNA SOLUCIÓN SENTIMENTALISTA AL "PROBLEMA MORAL"

TESIS QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN FILOSOFÍA
PRESENTA:

EMILIANO VILLARREAL ROSALES

TUTOR:
DR. SANTIAGO ECHEVERRI SALDARRIAGA
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS (IIF) UNAM

COMITÉ TUTOR:
DR. PAU LUQUE
DR. MIGUEL ÁNGEL SEBASTIÁN GONZÁLEZ
DR. MOISÉS VACA PANIAGUA
DR. RODRIGO VALENCIA PACHECO

CIUDAD DE MÉXICO, AGOSTO 2026

AGRADECIMIENTOS

Este trabajo solo ha sido posible gracias al apoyo de diferentes personas que me han acompañado a lo largo de mis estudios de maestría. Una de esas personas ha sido mi tutor, el Dr. Santiago Echeverri, quien en un momento incierto me ofreció dirigir mi proyecto de tesis y quien me ha introducido al ámbito profesional de la filosofía. Gracias, Santiago, por todo el trabajo que has puesto en mi trabajo. Mis padres, Héctor y Cecilia, han contribuido no solo con sus muestras permanentes de amor incondicional, sino también invitándome, desde siempre, a hacer lo que me gusta. A Elisa y a Lupita les agradezco por haber sido la casa a la que siempre he podido llegar. Los amigos que han formado parte de esta experiencia son muchos, gracias a todos ellos por compartir conmigo lo compartido. Quiero mencionar especialmente a Adi, a Julio y a San, porque lo compartido con ellos ha sido cotidiano. También quiero agradecer a Lola por leer lo que nadie había leído, al Dr. Eduardo García Ramírez por su guía durante el proceso de ingreso a la maestría y a los cuatro investigadores que han aceptado formar parte del jurado de mi examen de grado. Finalmente, quisiera agradecer a Julia por haber recorrido a mi lado una parte importante del camino.

Para la realización del trabajo contó con el apoyo de una beca otorgada por la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI). Asimismo, también se recibió financiamiento de parte de los proyectos de investigación “Subdeterminación y justificación” (SECIHTI IH-2025-I-126) y “Escepticismo y racionalidad” (PAPIIT IA 400124), ambos dirigidos por el Dr. Santiago Echeverri Saldarriaga.

ÍNDICE

Introducción general ……..…… 1

Capítulo I. El "problema moral" y la solución racionalista de Smith …….. 6

1. Introducción ……. 6
2. Las proposiciones que conforman el problema moral ……..8
   2.1. La objetividad del juicio moral ……..8
   2.2. El carácter práctico del juicio moral …… 10
   2.3. La "imagen estándar" de la psicología de la motivación para actuar …….. 16
3. Tipos de soluciones al problema moral …….. 19
   3.1. Soluciones parciales en la literatura …… 20
4. La solución racionalista de Smith …….. 23
5. Conclusiones …….. 27

Capítulo II. Objeciones a la solución racionalista de Smith al problema moral ……29

1. Introducción ……... 29
2. Un caso claro de juicio moral …… 29
3. Los elementos de la teoría racionalista de Smith …….. 31
4. Objeciones a la teoría racionalista …… 32
   4.1. Objeción de hiper-intelectualización ……32
   4.2. Objeción de justificación contingente …… 35
   4.3. Objeción de modo de presentación inadecuado …… 38
5. Conclusiones …… 42

Capítulo III. Una solución sentimentalista al problema moral ……. 41

1. Introducción …… 41
2. Concepción general de las emociones …….. 44
   2.1. Intencionalidad …….. 45
   2.2. Condiciones de corrección ……..46
   2.3. Diferencia respecto de los juicios ……...47
   2.4. Motivación de la acción …… 48
   2.5. Reflejos de preocupaciones ……49
3. Emociones morales ……... 51
4. Las emociones morales como bases de los juicios morales …… 52
5. El sentimentalismo moral como solución al problema moral …… 54
   5.1. Sentimentalismo y cognitivismo moral …… 54
   5.2. Sentimentalismo e internismo moral …… 56
   5.3. Sentimentalismo y la teoría humeana de la motivación …… 60
6. Lecciones derivadas de las objeciones al racionalismo de Smith …… 62
   6.1. Sentimentalismo y plausibilidad psicológica ……..63
   6.2. Sentimentalismo y cambio en la justificación ……65
   6.3. Sentimentalismo y modo de presentación actual ……67
7. Conclusiones …… 69

Conclusiones generales ……..... 70
Bibliografía …… 72

INTRODUCCIÓN

En el curso de la vida cotidiana los seres humanos acostumbramos emitir juicios morales. Tanto en el pensamiento como en el lenguaje, parte de nuestra experiencia consiste en evaluar moralmente correcta o incorrecta la propia conducta y la de otras personas. Estas evaluaciones determinan cómo nos desplegamos dentro del entorno social que nos rodea. Pues es con base en ellas que tomamos decisiones acerca de cuestiones tales como qué hacer y qué no hacer o con quiénes sí y con quienes no relacionarnos. Sin embargo, pese a la recurrencia con la que los llevamos a cabo y a lo influyentes que son en la configuración de nuestro contexto, articular una imagen teórica de los juicios morales no es una labor sencilla. En particular, cuando reflexionamos acerca del tipo de rasgos que caracterizan a estos juicios en el uso ordinario que les damos, nos percatamos de que algunos de esos rasgos parecen sostener una cierta tensión entre sí. Dos rasgos cuya atribución simultánea a los juicios morales se ha acostumbrado identificar como problemática son: la capacidad de desempeñar una función descriptiva y la capacidad de motivar la acción.

Cuando emitimos un juicio moral, parece apropiado pensar que describimos un hecho concreto. Por ejemplo, cuando juzgo moralmente correcto regresarle a un amigo dinero que me ha prestado, parece que lo que hago es describir el hecho concreto de que es moralmente correcto regresarle a mi amigo el dinero que me prestó. Y de manera análoga, cuando juzgo moralmente incorrecta la mentira de un político, parece que lo que hago es describir el hecho concreto de que la mentira del político es moralmente incorrecta. Más allá de que intuitivamente parece adecuado, atribuir una función descriptiva a los juicios morales permite explicar diferentes características del uso que le damos a estos juicios. En concreto, permite explicar por qué suponemos que un juicio moral puede ser verdadero o falso, por qué tendemos a pensar que los desacuerdos sobre cuestiones morales son desacuerdos genuinos o por qué la estructura gramatical de los juicios morales expresados en el lenguaje es la correspondiente a los enunciados descriptivos en general. Sin la atribución de una función descriptiva a los juicios morales, difícilmente podríamos dar cuenta de por qué características como las recién mencionadas se manifiestan en la forma en la que utilizamos estos juicios.

Por otro lado, la emisión de un juicio moral parece disponernos a actuar de ciertas maneras, específicamente, de maneras que sean congruentes con el juicio emitido. Por ejemplo, al juzgar moralmente correcto el aborto, parece que me dispongo a hacer cosas como promover el acceso al aborto, discutir un con un anti aborto o cuando menos no

impedir que una persona pueda abortar. Y de forma similar, si juzgo que sería moralmente incorrecto lastimar a mi gato, parece que eso me dispone a evitar cualquier acción que de algún modo pueda implicar dañar a mi gato. Aunque las disposiciones a actuar derivadas de la emisión de juicios morales no siempre determinan qué es lo que de hecho terminamos haciendo, sí parecen ejercer una influencia importante en la determinación del tipo de acciones que tenemos motivación de llevar a cabo. Prueba de lo anterior es que en general, suponer que existen estas disposiciones permite predecir con relativa confianza qué tipo de conductas son las que cabría esperar de una persona. Sin la atribución de una función práctica —en el sentido de que pueden llevar a la acción— a los juicios morales, difícilmente podríamos dar cuenta de por qué la elaboración de semejantes predicciones forma parte del uso que le damos a estos juicios.

La razón por la cual se tiende a pensar que es problemático atribuir a los juicios morales tanto una función descriptiva como una función práctica tiene su fundamento en una cierta concepción de la naturaleza de los estados psicológicos que pueden explicar cada una de esas funciones. De acuerdo con esa concepción, originalmente planteada por David Hume (1978/1739), las respectivas explicaciones de cada una de las funciones relevantes refieren a tipos de estados necesariamente diferentes. Por un lado, la función descriptiva se vincula con estados como las creencias que se distinguen por representar cómo es el mundo. Y por otro lado, la función práctica se vincula con estados como los deseos que se distinguen por representar cómo debería de ser el mundo. El aspecto problemático derivado de esta concepción es que, si cada función se explica por un tipo de estado diferente, entonces no es claro qué tipo de estado podría explicar un juicio moral. Pues mientras que si optamos por explicarlos en términos de estados como las creencias parece que no podríamos atribuirles una función práctica, si optamos por explicarlos en términos de estados como los deseos parece que no podríamos atribuirles una función descriptiva. Es así como llegamos a la pregunta de investigación a la que este trabajo intenta responder: ¿qué imagen de la psicología detrás de los juicios morales puede acomodar el hecho de que estos juicios desempeñan tanto una función descriptiva como una función práctica?

Una respuesta influyente a la pregunta anterior es desarrollada por Micheal Smith, en su libro “El problema moral” (1994). De acuerdo con Smith, para acomodar el hecho de que los juicios morales tienen tanto una función descriptiva como una función práctica, deberíamos de explicar estos juicios en términos de un determinado tipo de creencias. En concreto, Smith piensa que deberíamos de explicarlos en términos de creencias acerca de qué

es lo que un hipotético agente plenamente racional desearía que hiciéramos. Al adoptar este enfoque, la posibilidad de atribuir una función descriptiva a los juicios morales queda asegurada simplemente porque el tipo de estado que los explica es el correspondiente a una función descriptiva en general. Pero además, en tanto las creencias que explican a los juicios morales reflejan aquello que sería deseable desde una perspectiva de plena racionalidad, la posibilidad de atribuir una función práctica a estos juicios también parece quedar garantizada. Pues en la medida en que somos animales racionales, tiene sentido afirmar que aquello que creemos que sería lo más razonable que podríamos hacer muchas veces es lo que de hecho terminamos haciendo.

Pese a ser capaz de acomodar la atribución a los juicios morales de las dos funciones que consideramos apropiado atribuirles, la propuesta de Smith se articula como una teoría problemática por razones que dasarrollaré a lo largo del trabajo. En primer lugar, dada la complejidad del contenido de las creencias que supone que explican a los juicios morales, la teoría implica que para emitir un juicio moral un sujeto tendría que satisfacer requisitos cognitivos implausibles. En segundo lugar, al suponer que las creencias que explican a los juicios morales capturan veredictos fundados en la plena racionalidad, la teoría no puede dar cuenta del hecho de que la justificación de conductas específicas que ofrecen los juicios morales es susceptible al cambio, en el sentido de que puede incrementarse, disminuirse o incluso neutralizarse. Y en tercer lugar, al afirmar que los juicios morales de un sujeto lo motivan a actuar solo en la medida en que ese sujeto quiere actuar como un agente plenamente racional quisiera que lo haga, la teoría no puede dar cuenta del hecho de que, por sí sola, la emisión de un juicio moral puede motivar.

Partiendo de las tres problemáticas descritas en el párrafo anterior, la hipótesis que pondré a prueba a lo largo de este trabajo es que la teoría de Smith no constituye una respuesta satisfactoria a la pregunta de investigación. Sobre la base de ese diagnóstico, procederé a desarrollar una respuesta alternativa. De acuerdo con dicha respuesta alternativa, para acomodar el hecho de que los juicios morales tienen una función descriptiva y una función práctica, hay que explicar a estos juicios en términos de lo que Haidt (2003) denomina “emociones morales”. Como las emociones en general, las emociones morales son estados psicológicos que poseen múltiples componentes. Entre esos diferentes componentes, se encuentran tanto el tipo de rasgos que permiten explicar una función descriptiva, como el tipo de rasgos que permiten explicar una función práctica. Esto no significa que las emociones morales por sí solas puedan realizar todo el trabajo explicativo requerido para dar

cuenta de los juicios morales. Una explicación completa del tipo de funciones (descriptiva y práctica) de estos juicios no dejará de requerir de la referencia a los dos tipos de estados, necesariamente diferentes, que en un inicio se consideraban pertinentes para explicar dichas funciones. Sin embargo, en la medida en que la explicación de los juicios morales ya no recae únicamente sobre un único estado, la tensión que originalmente invitaba a pensar en la existencia de un problema queda resuelta. Pero además de acomodar las dos atribuciones de funciones que inicialmente queríamos acomodar, la explicación de los juicios morales en términos de emociones morales supone otros beneficios teóricos importantes. El primer beneficio es que una explicación de este tipo no genera los problemas que sí generaba la solución planteada por Smith. Y el segundo beneficio es que la explicación no involucra la reducción del ámbito de la moralidad al ámbito de la racionalidad. Esto último constituye un beneficio en tanto permite entender por qué muchas veces lo mejor en términos morales no es lo más racional y, además, porque permite ofrecer una imagen más rica del tipo de sensibilidad que fundamenta la emisión de juicios morales.

Este trabajo se compone de tres capítulos. El capítulo I tiene dos objetivos diferentes. Un primer objetivo es presentar detalladamente el problema relativo a la atribución a los juicios morales tanto de una función descriptiva como de una función práctica. Para ello, me remitiré a una formulación especialmente influyente de este problema, la cual es elaborada por el mismo Smith (1994) bajo la etiqueta de "problema moral". El problema moral se articula bajo la forma de un trilema. Dicho trilema se conforma de tres proposiciones aparentemente plausibles pero conceptualmente inconsistentes entre sí. Por otro lado, el segundo objetivo del capítulo I es llevar a cabo una reconstrucción de la solución al problema moral que Smith desarrolla. Como mencioné previamente, dicha solución se resume en la idea de que los estados psicológicos que explican a los juicios morales son creencias acerca de lo que un agente plenamente racional desearía que hiciéramos. Dada su fundamentación en la racionalidad, esta propuesta puede entenderse como una teoría racionalista de los juicios morales.

Por su parte, al igual que el capítulo I, el capítulo II de este trabajo también cumple con dos objetivos. El primer objetivo consiste en llevar a cabo dos clarificaciones diferentes. Mientras que la primera clarificación concierne a qué es exactamente lo que se entiende por juicio moral, la segunda apunta a dilucidar cuáles son los elementos que estructuran la explicación racionalista de estos juicios. El segundo objetivo del capítulo es desarrollar, bajo

la forma de objeciones, los tres problemas antes descritos que enfrenta la teoría racionalista de Smith.

Finalmente, el capítulo III tiene el único objetivo de desarrollar la propuesta de solución al problema moral que identifico como alternativa a la solución racionalista planteada por Smith. Dada su fundamentación en un cierto tipo de estados emocionales (las emociones morales), esta propuesta puede entenderse como una teoría sentimentalista de los juicios morales. Para el desarrollo de la propuesta, comenzaré caracterizando a las emociones en general. Después, pasaré a introducir el concepto de emociones morales. Más adelante, mostraré cómo una teoría semejante efectivamente puede explicar y acomodar las tres proposiciones que conforman al problema moral. Y por último, mostraré por qué la teoría sentimentalista que propongo no enfrenta ninguna de las objeciones que afecta a la teoría racionalista de Smith.

CAPÍTULO I

El "problema moral" y la solución racionalista de Smith

1. Introducción

Smith (1994) sostiene que, los principales desacuerdos que han direccionado el desarrollo de la discusión en metaética, pueden entenderse como desacuerdos derivados de la inconsistencia entre tres enfoques teóricos aparentemente plausibles: el cognitivismo moral, el internismo moral y la teoría humeana de la motivación. Esta inconsistencia es presentada por Smith bajo la forma de un trilema, al cual denomina el "problema moral" (PM). El trilema planteado por Smith (1994, p.12) es el siguiente:

> ***Problema moral (PM)*:**
> **i)** Los juicios morales de la forma "es correcto que yo haga φ" expresan las creencias de una persona sobre una cuestión objetiva de hecho, un hecho acerca de lo que es correcto hacer.
>
> **ii)** Si alguien juzga que es correcto hacer φ entonces, *ceteris paribus,* está motivado a hacer φ.
>
> **iii)** Un agente está motivado a actuar de cierto modo si tiene un deseo apropiado y una creencia medio-fin, donde creencias y deseos son, en la terminología de Hume, existencias distintas.

Las proposiciones (i) y (ii) refieren a rasgos que parecen caracterizar a los juicios morales en el contexto de la práctica moral ordinaria y cuya respectiva afirmación se articula como tesis central de una teoría metaética.[1] Mientras que (i) captura la tesis central del cognitivismo moral, (ii) recoge la tesis central del internismo moral. Por su parte, la proposición (iii) refiere a una cierta imagen respecto de la psicología de la motivación para actuar, específicamente, la imagen afirmada por la llamada "teoría humeana de la motivación".

[1]Aquí y en lo que sigue, al hablar de "práctica moral ordinaria" me estaré refiriendo al uso cotidiano que cualquier persona hace de juicios y conceptos morales en general. Esta etiqueta permite distinguir los usos cotidianos de juicios y conceptos morales de cualquier uso técnico o teórico.

La inconsistencia entre las tres proposiciones que conforman PM no es de naturaleza formal sino más bien conceptual (Kriegel, 2012, p. 469).[2] Se produce como consecuencia de que su afirmación simultánea implica afirmar dos conceptualizaciones aparentemente inconsistentes con respecto a la naturaleza psicológica de los juicios morales. En concreto, dos conceptualizaciones que difieren entre sí al especificar qué tipo de estados mentales son los expresados por juicios morales.[3] Según la proposición (i), los estados expresados por juicios morales son estados cognitivos como las creencias. Y de acuerdo con la proposición (ii), los estados expresados por juicios morales son estados que se relacionan intrínsecamente con la motivación para actuar. La inconsistencia entre estas dos conceptualizaciones se concreta una vez que se afirma la proposición (iii), la cual establece que los estados cognitivos y los estados intrínsecamente motivacionales, en su condición de tipos de estados mentales cuya posesión es requerida para la motivación para actuar, constituyen existencias distintas. Pues si (iii) está en lo cierto, entonces o bien (i) será verdadera y (ii) falsa, o bien (i) es falsa y (ii) verdadera. En cualquier caso, PM implicaría la falsedad de una idea que, desde la perspectiva de la práctica moral ordinaria, resultaba plausible.

El objetivo de este primer capítulo es llevar a cabo una reconstrucción tanto de PM como de la solución que Smith (1994) ofrece a dicho problema. El capítulo está estructurado de la siguiente manera. En la sección 2, expondré cuáles son las razones que respaldan la afirmación de cada una de las proposiciones que dan forma a PM. Más adelante, en la sección 3, pasaré a identificar tres posibles soluciones a PM que Smith ubica en la literatura en metaética. Y finalmente, en la sección 4, presentaré la propuesta de solución a PM desarrollada por el propio Smith.

[2] Kriegel (2012, p. 469) muestra cómo PM puede reformularse de manera tal que la inconsistencia en cuestión sí sea de naturaleza formal. De acuerdo con esa reformulación, el problema se configura como un trilema compuesto de las siguientes proposiciones: (1) Los juicios morales pretenden ser sobre cuestiones objetivas de hecho. (2) Los juicios morales son inherentemente motivacionales. (3) Los estados mentales, incluídos los juicios morales, no pueden tener tanto pretensión de objetividad como ser inherentemente motivacionales. Si bien en lo que sigue trabajaré sobre la formulación original de Smith (1994), tomar en cuenta la reformulación de Kriegel permite ganar claridad en la caracterización del problema.

[3] Aunque Smith no se pronuncia al respecto, la relación de expresión entre juicios morales y estados mentales puede entenderse como estableciendo que, cuando un sujeto realiza un juicio moral, dicho sujeto manifiesta su posesión de un estado mental, a saber, el estado mental expresado por su juicio. En tanto los juicios morales, como cualquier otro tipo de juicio, constituyen actos mentales, esta manifestación podrá permanecer en una dimensión privada o podrá adquirir una dimensión pública a través de su aseveración mediante un enunciado. Independientemente de si permanece en una dimensión privada o de si adquiere una dimensión pública, el aspecto relevante es que, con su realización de un juicio moral, un agente manifiesta su posesión de un cierto estado mental.

2. Las proposiciones que conforman el problema moral

A lo largo de esta sección me dedicaré a presentar cada una de las proposiciones que componen PM. Para ello, en lo particular de cada proposición haré dos cosas diferentes. Primero, comenzaré clarificando qué es lo que establece, y posteriormente, indicaré cuáles son las razones que, de acuerdo con Smith, respaldan su plausibilidad. Con esto, lo que me interesa es dar cuenta de porqué PM constituye un trilema genuino, conformado por distintas proposiciones a cuya verdad no deberíamos de estar dispuestos a renunciar.

2.1. La objetividad del juicio moral

La proposición (i) de PM establece que "los juicios morales de la forma "es correcto que yo haga p" expresan las creencias de una persona sobre una cuestión objetiva de hecho, un hecho acerca de lo que es correcto hacer" (Smith, 1994, p.12). Esta proposición captura un supuesto que opera en el contexto de la práctica moral ordinaria. De acuerdo con dicho supuesto, al que Smith (1994, p.6) denomina "la objetividad del juicio moral", los juicios morales desempeñan una función descriptiva acerca de hechos objetivos en el mundo. Y puesto que los estados mentales que desempeñan una función descriptiva acerca de hechos objetivos en el mundo son estados cognitivos, lo que (i) termina por afirmar es que los juicios morales expresan estados mentales cognitivos. Dicho en otras palabras, lo que (i) termina por afirmar es la verdad del cognitivismo moral.[4]

La noción de objetividad empleada en (i) busca señalar que, al menos en principio, los hechos descritos por juicios morales son igualmente accesibles para cualquier persona, independientemente de cuál sea su perspectiva particular. Para determinar la corrección moral de una acción, lo único que habrá que tomar en consideración son aspectos correspondientes a las circunstancias en las que la acción relevante ha tenido lugar. De ahí que, si por ejemplo, dos sujetos juzgan la corrección moral de una determinada acción, en la medida en que ambos tengan el mismo conocimiento acerca de las circunstancias en que la acción relevante

[4] El cognitivismo moral, entendido como un enfoque teórico cuya tesis central es que los juicios morales expresan estados mentales cognitivos, se articula como un compromiso inherente al realismo moral en general (tanto en sus variantes naturalistas como no-naturalistas). Sin embargo, su afirmación también es incorporada por ciertas formas de irrealismo moral, en particular, por la llamada "teoría del error" desarrollada originalmente por (Mackie, 1977).

ha tenido lugar, lo que cabrá esperar es que haya coincidencia entre sus respectivos juicios morales.

En su presentación de (i) como una de las proposiciones que componen PM, Smith no desarrolla ningún argumento positivo que respalde su plausibilidad.[5] Más bien, adopta una estrategia negativa, dedicada a explorar la viabilidad de una respuesta satisfactoria a lo que identifica como el "reto expresivista" en contra de la atribución de una función descriptiva a los juicios morales. Este reto, planteado inicialmente por Ayer (1936), se estructura bajo la forma de un dilema. De acuerdo con dicho dilema, si efectivamente los juicios morales desempeñan una función descriptiva, entonces, los hechos morales que describen solo pueden ser hechos naturales o hechos no-naturales. Sin embargo, puesto que resulta problemático concebir a los hechos morales tanto en términos de hechos naturales como en términos de hechos no-naturales, hay que concluir que los juicios morales en realidad no desempeñan una función descriptiva.[6] Frente al reto expresivista, Smith defiende la viabilidad de una cierta concepción naturalista de los hechos morales. Específicamente, afirma que a través del método del análisis conceptual, los hechos morales pueden entenderse como hechos acerca de razones para la acción.[7] Pero la afirmación de semejante posibilidad no contribuye de ninguna manera a comprender qué tipo de razones son las que hacen de (i) una proposición plausible.

De acuerdo con Kriegel (2012, p.470), la decisión de Smith de no hacer explícita ninguna razón positiva a favor de (i) responde al reconocimiento de esta proposición como un presupuesto "fenomenológicamente manifiesto" en la experiencia que un sujeto tiene cuando emite un juicio moral. Se trata, pues, de una proposición que refiere a un aspecto inherente al acto mismo de juzgar moralmente. En particular, la verdad de (i) parece darse por sentada en aquel ámbito de la práctica moral ordinaria que consiste en participar de conversaciones y discusiones que giran en torno a desacuerdos morales. Pues en estos contextos de desacuerdo,

[5] Cabe mencionar que la decisión de Smith de no presentar un argumento positivo a favor de (i) no refleja una falta de argumentos generalizada en la literatura en metaética. En metaética, uno de los argumentos más discutidos se fundamenta en lo que se conoce como el problema Frege-Geach. De acuerdo con este argumento, debemos de aceptar que los juicios morales expresan estados mentales cognitivos porque solo así podemos dotar de sentido a la capacidad de estos juicios para participar en estructuras inferenciales. Para una exposición detallada del problema Frege-Geach y del argumento a favor del cognitivismo moral que de él se desprende, véanse (Schroeder, 2007) y (Wood, 2017).

[6] Según Ayer (1936), concebir a los hechos morales en términos de hechos no-naturales es problemático porque no hay una explicación epistemológica plausible de cómo accedemos a esos hechos, y concebir a los hechos morales en términos de hechos naturales es problemático porque implica cometer una falacia naturalista.

[7] En este punto, Smith no muestra cómo se articularía una concepción semejante, sino que únicamente afirma su viabilidad y señala que la cuestión quedará clarificada hacia el final del libro, una vez que haya desplegado su teoría racionalista de los juicios morales.

(i) opera como aquello que dota de sentido a la práctica en cuestión. Al conversar sobre lo que sería moralmente correcto hacer frente a una determinada circunstancia, comunicamos nuestro propio juicio moral y escuchamos el de aquellos con quienes conversamos porque suponemos que el intercambio contribuirá a la formulación de un juicio moral verdadero, que describa adecuadamente un hecho moral; en concreto, el hecho moral concerniente a la corrección moral de la acción que evaluamos. Y de manera análoga, cuando discutimos con alguien sobre la corrección moral de una acción, presentamos razones en sustento de nuestro propio juicio moral y en contra del de aquellos con los que discutimos porque suponemos que ello contribuirá a mostrar la verdad de nuestro juicio, es decir, contribuirá a mostrar que nuestro juicio describe adecuadamente un hecho moral. Es precisamente en la medida en que el cognitivismo moral permite dotar de sentido a aspectos centrales de nuestra práctica moral ordinaria—como aquellos que giran en torno al desacuerdo moral— que una concepción adecuada de los juicios morales debería de poder afirmar la verdad de (i).

2.2. El carácter práctico del juicio moral

La proposición (ii) de PM señala que "si alguien juzga que es correcto hacer φ, entonces, *ceteris paribus,* está motivado a hacer φ" (Smith, 1994, p.12) . De acuerdo con Smith, al igual que la proposición (i), esta segunda proposición refiere también a un rasgo que caracteriza a los juicios morales en el contexto de la práctica moral ordinaria. Este rasgo, al que Smith (1994, p.7) llama "el carácter práctico del juicio moral", captura el hecho de que del juzgar moralmente suele seguirse la obtención de un cierto tipo de motivación para actuar. Específicamente, motivación para actuar de formas que se ajustan al sentido de la corrección moral marcada por el juicio relevante. Por ejemplo, si juzgo moralmente incorrecta la acción de mi vecino de golpear a su esposa, entonces, en condiciones normales tendré motivación para llevar a cabo acciones que impidan que lo siga haciendo. Estaré motivado, pues, a realizar acciones como ir a confrontar a mi vecino o llamar a la policía. De acuerdo con Smith, la idea plasmada en (ii) puede entenderse en términos de una cierta formulación de la tesis central del internismo moral.

En metaética, el internismo moral constituye un enfoque cuya tesis central establece que existe una conexión entre juicios morales y motivación para actuar cuyo fundamento se ubica en el interior de los juicios morales, es decir, en su contenido. Sin embargo, si partiendo de esta afirmación se derivara un compromiso ulterior con la tesis de que "[...] el juicio moral

trae consigo motivación *simpliciter*" (Smith, 1994, p.61), entonces el internismo se articularía como una teoría falsa. Esto se debe a la existencia de casos de sujetos como los psicópatas, quienes pese a juzgar moralmente no encuentran motivación para actuar. Dado su reconocimiento de la existencia de este tipo de contraejemplos a la tesis internista tomada sin más, Smith introduce un elemento que le permite acotar el alcance del internismo afirmado en (ii) de manera que obtenga inmunidad a semejantes contraejemplos. El elemento en cuestión es el operador *ceteris paribus*. Con la introducción de este operador, Smith especifica que el internismo en (ii) es un internismo *débil*. Esto significa que la conexión que contempla entre juicio moral y motivación no se cuantifica universalmente, sino que se restringe a ciertos casos en particular. En concreto, se restringe a lo particular del tipo de sujetos que no podrían protagonizar casos que se articulan como contraejemplos a la tesis internista. El criterio que Smith utiliza para indicar qué es lo que distingue a ese tipo de sujetos es la ausencia de irracionalidad práctica. Pues desde su perspectiva, es precisamente la irracionalidad práctica lo que explica por qué sujetos como los psicópatas no obtienen motivación al emitir un juicio moral.

En su desarrollo del tipo de internismo contemplado en la proposición (ii), Smith (1994, pp. 61-62) caracteriza dicho enfoque atribuyéndole un compromiso con las siguientes dos tesis:

> ***Tesis de irracionalidad (TI)*:** Si un sujeto *S* juzga moralmente correcto hacer *φ* en ciertas circunstancias *C*, entonces, o bien *S* está motivado a hacer *φ* en *C*, o bien *S* es irracional en términos prácticos.[8]
>
> ***Tesis de posesión de razón (TPR):*** Si un sujeto *S* juzga moralmente correcto hacer *φ* en ciertas circunstancias *C*, entonces, *S* tiene una razón para hacer *φ* en *C*.[9]

Estas dos tesis guardan una estrecha relación entre sí. En la medida en que *TPR* parece implicar *TI*, *TPR* puede tomarse como una explicación de *TI*. Pues si juzgar moralmente correcta una acción consiste en tener una razón para llevar a cabo esa acción, entonces, es razonable concluir que, si un sujeto no está motivado a actuar como juzga moralmente correcto hacerlo, al no estar motivado a actuar como cree que tiene razón para hacerlo, dicho

[8] La atribución al internismo moral de un compromiso con esta tesis puede encontrarse en (Blackburn, 1984), (Johnston, 1989) y (Pettit y Smith, 1993).
[9] La identificación de esta tesis como uno de los compromisos propios del internismo moral se da tanto en (Nagel, 1970) como en (Korsgaard, 1986).

sujeto es irracional.[10] La falta de motivación para hacer lo que se considera moralmente correcto implica que hay un sentido en el cual no se está siendo sensible a las razones disponibles y, en consecuencia, no se está cumpliendo con lo que cabría esperar de un sujeto racional en general. En resumen, *TI* y *TPR* señalan que cuando un sujeto no encuentra motivación para actuar en conformidad con su juicio moral, manifiesta aquel tipo de irracionalidad práctica a la que comúnmente se denomina debilidad de la voluntad o akrasia, la cual consiste en ir en contra del mejor juicio propio acerca de lo que se debería de hacer. Pues en cualquier caso, el mejor juicio de un sujeto será aquel que cuente con el apoyo de las razones que están a su disposición .

A través del reconocimiento de *TI* y *TPR* como dos tesis que perfilan el tipo de internismo referido en (ii), Smith clarifica porqué los juicios morales no traen consigo motivación *simpliciter*. En concreto, la razón por la cual esto es así es porque los juicios morales sólo motivan si se satisface una condición de racionalidad práctica. Si un sujeto juzga moralmente y es racional en términos prácticos, entonces, encontrará motivación para actuar de formas que se ajusten con ese acto de juzgar.[11] Pero si un sujeto juzga moralmente y es irracional en términos prácticos, entonces, dicho sujeto no encontrará ninguna motivación. La primera afirmación tiene su fundamento en *TPR*, pues es ahí donde se reconoce que si hay racionalidad práctica, el vínculo entre juicio moral y motivación para actuar estará asegurado. Y la segunda afirmación se cimienta sobre *TI*, donde se indica que, en ausencia de racionalidad práctica, nada garantiza que el vínculo entre juicio moral y motivación para actuar efectivamente vaya a concretarse. Así, el internismo moral contenido en la proposición (ii) de PM puede formularse de la siguiente manera:

> ***Internismo moral débil racionalista (IMDR):*** en sujetos racionales en términos prácticos, existe una relación entre juicio moral y motivación para actuar que tiene un fundamento interno a los juicios morales.

[10] El tipo de razones para la acción del que se habla en TPR son razones normativas. Más adelante en este capítulo, durante el desarrollo de la proposición (iii) de PM, quedará clarificado en qué consisten este tipo de razones.

[11] Es importante subrayar la distinción entre la motivación para actuar y el actuar. La afirmación de necesidad internista vincula al juzgar moralmente con la obtención de un cierto tipo de motivación para actuar, a saber, motivación para actuar de formas que se ajusten al contenido del juicio relevante. Por lo anterior, la verdad de dicha afirmación de necesidad no depende de que todos los agentes respecto de los cuales se predica dicha afirmación actúen de una manera en particular, sino únicamente de que su juzgar moralmente se refleje en alguna medida en su motivación para llevar a cabo acciones que se ajustan a su juicio.

Tras la especificación del tipo de internismo moral asociado con la proposición (ii) de PM, Smith pasa a desarrollar un argumento que respalde la plausibilidad de una teoría semejante. Se trata de un argumento a la mejor explicación que tiene su punto de partida en la identificación de un cierto "hecho sorprendente" (Smith, 1994, p. 101). El hecho sorprendente en cuestión es que, en sujetos que no padecen ninguna forma de irracionalidad práctica, "[...] un *cambio en la motivación* se produce confiablemente cuando se produce un *cambio en el juicio moral* [...]" (Smith, 1994, p. 101).

Para ilustrar el hecho referido por Smith, imaginemos una discusión protagonizada por dos sujetos, Frida y Andrea. En el marco de esta discusión, Frida presenta a Andrea una serie de argumentos por los cuales juzga moralmente incorrecto comer carne. Tras escuchar los argumentos de Frida, Andrea, quien inicialmente juzgaba moralmente correcta esa misma acción, reconoce explícitamente haberse convencido de que su interlocutora tiene razón y afirma que su juicio moral ha cambiado, de manera que ahora juzga incorrecto el consumo de carne. Dada esta serie de eventos, el hecho sorprendente referido por Smith podrá reconocerse una vez que se plantea la pregunta de qué es lo que cabría esperar de la motivación para actuar de Andrea después del cambio en su juicio moral. Pues ante esta interrogante, lo más razonable sería afirmar que lo que cabrá esperar es que Andrea haya adquirido motivación para no comer carne. Es decir, que como consecuencia de la revisión de su juicio moral original, la motivación para actuar de Andrea se haya visto modificada en alguna medida en un sentido que se ajusta a su nuevo juicio moral revisado. Esto no significa que necesariamente Andrea ya no encontrará motivación para comer carne, sino únicamente que en alguna medida su motivación original para hacerlo habrá cambiado.

Una vez reconocido el hecho sorprendente que vincula el cambio en el juicio moral con el cambio en la motivación para actuar, el siguiente paso en el argumento de Smith consiste en distinguir entre dos explicaciones diferentes acerca de ese hecho: la internista y la externista. En apego a la caracterización del internismo moral elaborada al inicio de la presente sección, la explicación internista se distingue por afirmar que, del cambio en el juicio moral se sigue el cambio en la motivación para actuar, porque los juicios morales sostienen una relación intrínseca con la motivación para actuar. Desde la perspectiva internista, el contenido de los juicios morales es de una naturaleza tal que, al menos en lo particular de los sujetos racionales en términos prácticos, implica necesariamente motivación para actuar. Es por ello que después del cambio en el juicio moral de Andrea, lo que cabe esperar es que también haya un cambio en su motivación para actuar. Por su parte, la explicación externista del mencionado hecho sorprendente se centra en la apelación a factores

externos a los juicios morales. Por ejemplo, una explicación externista típica consistiría en atribuir a un sujeto como Andrea algún tipo de predisposición psicológica que le lleva a ajustar su motivación para actuar con su juicio moral. Así, desde el externismo, el hecho de que del cambio en el juicio moral se siga un cambio en la motivación no es el reflejo de una relación necesaria sino meramente contingente.

Llegado a este punto, el argumento a la mejor explicación de Smith pasa a una etapa final dedicada a mostrar que la explicación internista del hecho sorprendente es mejor que la explicación externista. Para mostrar que esto es así, Smith argumenta que la explicación externista es problemática en tanto sólo puede dar cuenta del hecho sorprendente entendiéndolo como una manifestación de fetichismo moral. En general, la palabra "fetichismo" designa un tipo de atribución de valor que no se fundamenta en el reconocimiento de propiedades inherentes al objeto valorado, sino en la mediación de una abstracción que hace que ese objeto parezca valioso. Las valoraciones fetichistas pueden atribuirse a objetos tales como modos de actuar. Por ejemplo, si un sujeto considera valioso comprar un reloj porque piensa que le dará prestigio social y no por alguna de las cualidades que de hecho distinguen a ese reloj en particular, entonces, el modo de actuar de dicho sujeto será fetichista. Y de la misma manera, si un sujeto considera valioso comenzar una relación sentimental con una persona de un país diferente al suyo porque siempre ha fantaseado con estar con alguien de ese país y no por cómo es la persona con la que se relacionará, entonces, el modo de actuar de dicho sujeto también será fetichista. En ambos casos, la mediación de una abstracción es lo que dota de valor al modo de actuar en cuestión. Para el comprador del reloj, la abstracción relevante dota de valor a la compra del reloj porque vincula ese modo de actuar con la adquisición de prestigio social. Y para el sujeto en el segundo caso, la abstracción relevante dota de valor a la decisión de comenzar una relación porque vincula ese modo de actuar con la satisfacción de una fantasía.

El fetichismo moral referido por Smith en el desarrollo de su argumento a la mejor explicación en contra del externismo puede entenderse como una aplicación del concepto general de fetichismo a lo particular de las atribuciones de valor moral. En concreto, el fetichismo moral tiene lugar en el contexto de atribuciones de valor moral a modos de actuar específicos. De acuerdo con Smith, la atribución de valor moral a un modo de actuar es fetichista cuando esa atribución se fundamenta en la mediación de una abstracción que presenta al modo de actuar en cuestión como *lo moralmente correcto*.[12] Por ejemplo,

[12] La expresión "lo moralmente correcto" funciona como etiqueta genérica para referir a cualquier descripción de la corrección moral como categoría abstracta. Sin embargo, el fetichismo moral puede articularse a partir de

pensemos en el caso de un sujeto que, después de percatarse de que una persona ha tirado accidentalmente su cartera en la calle, decide recoger la cartera y entregársela a su dueño. Si lo que explica por qué el sujeto entregó la cartera es una creencia de que semejante modo de actuar es *lo moralmente correcto*, entonces, según lo indicado por Smith, el caso contará como una manifestación de fetichismo moral. En contraste, si lo que explica la decisión del sujeto de entregar la cartera es algo como la preocupación por el bienestar de la persona que la tiró o una tendencia a la honestidad, entonces, el caso no presentará una manifestación de fetichismo moral. La diferencia entre el caso fetichista y el caso no-fetichista es planteada por Smith (1994, p. 73-74) en términos de la distinción entre contenido *de dicto* y contenido *de re*. En el caso fetichista, la atribución de valor moral se da a través de un estado mental que tiene como contenido una descripción *de dicto* respecto de la corrección moral. En cambio, en el caso no fetichista, la atribución de valor moral que motiva al sujeto a hacer lo que hace se da a través de un estado mental *de re* respecto de la corrección moral.

Una vez clarificada la cuestión de qué es lo que se entiende por fetichismo moral, puede apreciarse a qué es a lo que apunta el argumento a la mejor explicación de Smith en contra del externismo. El argumento afirma que, la explicación externista del hecho sorprendente es problemática porque únicamente puede dar cuenta de ese hecho entendiéndolo como una manifestación de fetichismo moral. Recordando la caracterización general de la explicación externista, de acuerdo con esa perspectiva, la conexión entre juicio moral y motivación tiene un fundamento externo a los juicios morales. Por lo regular, el externista afirma que el fundamento de dicha conexión se ubica en factores psicológicos que median entre el sujeto que juzga y el contenido de su juicio. Pero si esto es así, entonces, puesto que la motivación que produce un juicio moral nunca se fundamenta directamente en el contenido del juicio, sino más bien en una abstracción acerca del contenido del juicio, al dar cuenta de algo como el mencionado hecho sorprendente, el externista no tendrá más opción que afirmar el fetichismo moral. En contraste, la explicación internista, al sostener que el fundamento de la relación entre juicio moral y motivación se ubica en el contenido mismo de los juicios, no corre riesgo de afirmar el fetichismo.

---

cualquier otra expresión equivalente que describa a la corrección moral como categoría abstracta. Por ejemplo, desde una perspectiva utilitarista, la expresión correspondiente sería una como “lo que maximiza el bienestar de la mayoría”. Y desde una perspectiva deontológica, la expresión relevante sería una como “lo que querríamos que se reconociera como máxima universal”.

2.3. La "imagen estándar" de la psicología de la motivación para actuar

Finalmente, la proposición (iii) de PM establece que "un agente está motivado a actuar de cierto modo si tiene un deseo apropiado y una creencia medio-fin, donde creencias y deseos son, en la terminología de Hume, existencias distintas" (Smith, 1994, p.12). A diferencia de las proposiciones (i) y (ii), esta proposición no captura un rasgo atribuido a los juicios morales desde la práctica moral ordinaria, sino que refiere más bien a una cierta concepción de la psicología subyacente a la motivación para actuar. Dicha concepción, a la que Smith (1994, p. 7) denomina "la imagen estándar de la motivación para actuar", corresponde a lo que en metaética y filosofía de la acción se conoce como "teoría humeana de la motivación". De acuerdo con esta teoría, desarrollada originalmente por David Hume (1978/1739), creencias y deseos son dos tipos de estados mentales necesariamente diferentes, ambos requeridos para la explicación de cualquier acción intencional.

En su caracterización de los compromisos inherentes a la teoría humeana de la motivación (THM), Smith (1994, p. 92-93) distingue entre dos tesis diferentes, una más fuerte que la otra. La tesis fuerte establece que "la motivación tiene su fuente en la presencia de un deseo relevante y una creencia medio-fin" (Smith, 1994, p. 92). Y por otra parte, la tesis débil contempla que "la motivación requiere de la presencia de un deseo relevante y una creencia medio-fin" (Smith, 1994, p.93). El contraste en términos de fuerza entre estas dos tesis es consecuencia de que mientras una de ellas, la fuerte, se articula como una afirmación de necesidad, la otra, la débil, constituye más bien una afirmación de suficiencia. La la tesis fuerte afirma que la presencia de un deseo y una creencia medio-fin es una condición *necesaria* para que haya motivación. En cambio, la tesis débil afirma que la presencia de un deseo y una creencia medio-fin es una condición *suficiente* para que haya motivación.

Desde la perspectiva de Smith, las dos tesis con las que THM se compromete son verdaderas. En particular, considera que la tesis fuerte captura "una verdad simple pero importante" acerca de la naturaleza de la motivación (Smith, 1994, p.93). La verdad en cuestión es que tener motivación consiste siempre en tener un objetivo. Para Smith, esta verdad podrá reconocerse siempre que se cumplan tres condiciones. Primero, que se reconozca adecuadamente acerca de qué tipo de razones para la acción versa THM. Segundo, que se posea una concepción adecuada de la actitud de *desear*. Y tercero, que se comprenda correctamente qué tipo de explicación de la acción intencional es la que puede fundarse en el tipo de razones para la acción acerca de las cuales versa THM. En lo que sigue, me dedicaré a aclarar cada una de las tres condiciones recién mencionadas, pues al entender qué es lo que

respalda a cada una de esas condiciones, puede entenderse también por qué, según Smith, THM es una teoría plausible.

De acuerdo con Smith, para percatarse de la verdad de THM, hay que comenzar por reconocer de qué habla THM. Para ello, resulta imprescindible remitirse a la distinción entre dos tipos diferentes de razones para la acción: las razones normativas y las razones motivacionales. La afirmación "*S* tiene una razón para hacer *φ*" es ambigua en la medida en que puede tener dos significados diferentes. Por un lado, puede interpretarse en términos de razones normativas. Interpretada de esta manera, lo que la afirmación significa es que *S* tiene una justificación para hacer *φ*. Por otro lado, la afirmación puede interpretarse en términos de razones motivacionales. Interpretada de esta manera, lo que significa es que *S* posee un estado psicológico que podría explicar, a manera de antecedente, su hacer *φ*. Tanto las razones normativas como las razones motivacionales tienen la capacidad de hacer inteligible una acción intencional. Sin embargo, la dimensión desde la cual lo hacen es distinta. Las razones normativas hacen inteligible a la acción intencional desde la dimensión de la justificación. En cambio, las razones motivacionales hacen inteligible a la acción intencional desde una dimensión explicativa.[13]

Para Smith, entender THM requiere de la comprensión de este contraste entre tipos de razones para la acción. Esto se debe a que THM es, única y exclusivamente, una teoría acerca de las razones motivacionales. Tanto la tesis débil como la tesis fuerte de THM son tesis con respecto a la naturaleza de las razones motivacionales. Partiendo del reconocimiento del hecho anterior, podemos reformular cada una de esas tesis. La tesis débil puede reformularse de manera que lo establezca es que, la presencia de un deseo y una creencia medio-fin es una condición *suficiente* para tener una razón motivacional. Y la tesis fuerte puede reformularse de modo que lo que afirme es que la presencia de un deseo y una creencia medio-fin es una condición *necesaria* para tener una razón motivacional.

Ahora bien, dado el protagonismo de los deseos en la configuración de THM, para entender de forma adecuada esta teoría habrá que asegurarse de tener una concepción adecuada de dichos estados mentales. De acuerdo con Smith, una concepción adecuada de los deseos se configura en dos pasos. Primero, un paso dedicado a reconocer que los deseos son estados

[13] Más adelante, en la sección 4 de este capítulo, al reconstruir la propuesta de solución a PM de Smith, se podrá apreciar que hay un sentido en el cual las razones normativas también pueden explicar a la acción intencional. Sin embargo, pese a esa capacidad explicativa, el rasgo más característico de este tipo de razones es la capacidad de justificar la acción.

mentales con contenido proposicional. Y segundo, un paso dedicado a especificar qué tipo de contenido proposicional es propio de los deseos.

Para mostrar que los deseos son estados mentales con contenido proposicional, Smith (1994, p.133) desarrolla una objeción en contra de la que identifica como la principal concepción de los deseos como estados carentes de contenido proposicional. Dicha concepción es la concepción fenomenológica del deseo (CFD). De acuerdo con esta concepción, "los deseos son, como las sensaciones, simple y esencialmente estados que tienen cierto contenido fenomenológico" (Smith, 1994, p.133). Desear no consiste en sostener una cierta relación con una proposición, sino más bien, en atravesar una experiencia consciente.[14] La objeción de Smith en contra de CFD establece que, en tanto desde la perspectiva del sentido común no pensamos que los deseos son estados que poseen cualidades fenomenológicas esenciales, podemos concluir que esta concepción de los deseos es falsa. Su comprobación de la ausencia de semejante componente fenomenológico en nuestra imagen de sentido común de los deseos se fundamenta en el planteamiento de múltiples escenarios hipotéticos. En primera instancia, Smith propone pensar en el caso del deseo de un padre de que le vaya bien a su hijo. Aunque la mayor parte del tiempo el hombre no es consciente de su posesión de un deseo semejante, esta intermitencia en la consciencia respecto a su posesión del deseo relevante no nos lleva a dudar del hecho de que efectivamente lo posee, ni tampoco a pensar que hay ocasiones en las que su compromiso con dicho deseo se encuentra suspendido. Por otro lado, también podemos encontrar casos relativos a experiencias en primera persona.[15] Por ejemplo, las experiencias de cruzar una calle o de escribir un texto. En este tipo de situaciones, el hecho de que no parezca haber ningún deseo del cual seamos conscientes, no es impedimento para que nos auto-atribuyamos un deseo de hacer lo que estamos haciendo. Con base en los escenarios recién descritos, Smith considera haber mostrado la falsedad de CFD y, por consecuencia, la verdad de la afirmación según la cual los deseos son estados con contenido proposicional.

Tras mostrar que los deseos tienen contenido proposicional, Smith pasa a centrar la atención en la especificación del tipo de contenido propio de estos estados. Para ello, se remite a la noción de "dirección de ajuste".[16] La noción de dirección de ajuste constituye una herramienta que permite distinguir entre dos tipos diferentes de estados mentales. De un lado

[14] De acuerdo con Smith (1994, p.105), este uso del término "sensaciones" proviene de (Hume, 1888, p.414).
[15] El planteamiento original de este tipo de casos proviene de Stroud (1977).
[16] Históricamente, la noción de dirección de ajuste tiene su origen en ciertas intuiciones teóricas planteadas por Anscombe (1957). Pero tanto la introducción de la etiqueta "dirección de ajuste" como el uso que Smith hace de la noción provienen de (Platts, 1979).

de la distinción se ubican estados que apuntan a la verdad, en el sentido de que buscan que su contenido se ajuste al mundo. A estos estados se les atribuye la dirección de ajuste *mente-a-mundo*. Y del otro lado de la distinción se ubican estados que apuntan a la realización, en el sentido de que buscan que el mundo se ajuste a su contenido. A estos estados se les atribuye la dirección de ajuste *mundo-a-mente*.

Desde la perspectiva de Smith, los deseos se configuran como un ejemplo paradigmático del tipo de estados mentales a los que se les atribuye la dirección de ajuste *mundo-a-mente*. Es decir, son estados que apuntan a la realización y no a la verdad. Más allá de la caracterización funcional que implica semejante concepción de los deseos, el reconocimiento de este hecho es importante porque es lo que, según Smith, permite entender el sentido en el cual creencias y deseos son "existencias distintas". Pues las creencias, en contraste con los deseos, se configuran como el ejemplo paradigmático del tipo de estados mentales con dirección de ajuste *mente-a-mundo*.

Una vez descrita la concepción de los deseos adecuada para comprender THM, Smith pasa a la etapa final de su exposición de esta teoría. Ahí, la atención recae sobre la especificación de la estructura propia de las explicaciones de la acción intencional fundadas en razones motivacionales. De acuerdo con Smith, dado el carácter de antecedente psicológico propio de las razones motivacionales, intuitivamente podríamos pensar que al explicar la acción φ de un sujeto *S* mediante la referencia a la razón motivacional R, lo que hacemos es afirmar que hay una relación causal entre la posesión de R, de parte de S, y su hacer φ. Sin embargo, independientemente de su plausibilidad inicial, Smith considera que esta idea comete el error de ignorar la perspectiva de *S*. A su parecer, R explica el hacer φ de *S* porque permite dar cuenta del hecho de que desde la perspectiva de *S*, hacer φ tuvo sentido. Así, Smith advierte que en realidad, la estructura propia de las explicaciones de la acción intencional en términos de razones motivacionales es una estructura teleológica. Pues solo si R captura un objetivo de *S*, R podrá dotar de sentido el hacer φ de *S*.

Llegado a este punto, Smith considera haber reunido los elementos suficientes para dar cuenta de por qué THM es una teoría que captura "una verdad simple pero importante acerca de la naturaleza de las razones motivacionales" (Smith, 1994, p.93). Al afirmar que tener una razón motivacional consiste en tener una creencia y un deseo, THM hace inteligible el hecho de que tener una razón motivacional y, por tanto, tener motivación, consiste en tener un objetivo.

3. Tipos de soluciones al problema moral

Como mencioné en la sección 1 de este capítulo, la incompatibilidad capturada por PM es de naturaleza conceptual. Específicamente, se trata de una incompatibilidad que concierne a la naturaleza conceptual de los estados mentales expresados por juicios morales. La cuestión puede plantearse de la siguiente manera. De acuerdo con el cognitivismo plasmado en la proposición (i), los juicios morales expresan estados cognitivos como las creencias. Y según el internismo descrito en la proposición (ii), los juicios morales expresan estados cuyo contenido se vincula necesariamente con la motivación para actuar. Sin embargo, la teoría humeana de la motivación referida en la proposición (iii) establece que ningún estado mental singular y unitario puede tener tanto el tipo de contenido que caracteriza conceptualmente a las creencias como el tipo de contenido que se vincula necesariamente con la motivación para actuar. Por lo tanto, para evitar afirmar una imagen conceptualmente contradictoria de los juicios morales, habrá que elegir entre renunciar al cognitivismo o al internismo o rechazar la teoría humeana de la motivación.

Sobre la base de la descripción anterior de la tensión que da pie a PM, podemos distinguir entre dos tipos diferentes de posibles soluciones al problema. Un primer tipo de solución, al que llamaré "solución parcial", consiste en negar alguna de las tres proposiciones que conforman PM. De acuerdo con Smith, hasta antes del desarrollo de su solución racionalista, en la literatura en metaética únicamente podían encontrarse este tipo de soluciones. Por otro lado, un segundo tipo de solución al que denominaré "solución conciliadora", consiste en mostrar, a través de la apelación a algún recurso teórico ulterior o de la reinterpretación de las proposiciones involucradas, que en realidad PM no constituye un trilema genuino, en tanto las proposiciones que lo conforman sí pueden afirmarse simultáneamente. Es este tipo de solución el que Smith considera haber logrado a través del desarrollo de su teoría racionalista de los juicios morales.

3.1. Soluciones parciales en la literatura

En la literatura en metaética, Smith (1994, pp. 126-129) ubica tres estrategias diferentes para resolver la tensión capturada en PM. Todas estas estrategias se configuran como soluciones parciales, en el sentido de que renuncian a alguna de las proposiciones del problema. Las estrategias en cuestión son la expresivista, la externista y la anti-humeana. A continuación,

expondré brevemente en qué consiste cada una de estas estrategias y qué es lo que Smith encuentra de problemático en ellas.[17]

La estrategia expresivista consiste en afirmar la verdad de las proposiciones (ii) y (iii) y renunciar a la de (i). El eje en el desarrollo de esta estrategia se ubica en la idea de que los juicios morales no expresan creencias representacionales ordinarias ni ningún otro tipo de estado cognitivo, sino más bien estados no-cognitivos como aprobación o desaprobación.[18] Para el expresivismo, un juicio o enunciado que afirma que una cierta acción es moralmente correcta o incorrecta no es apto para la evaluación en términos de criterios de verdad.[19] El argumento principal en defensa del expresivismo se remite a las dificultades inherentes a intentar dar cuenta tanto de la naturaleza metafísica de los hechos morales como del acceso epistémico que tenemos a ellos. Más específicamente, a lo que apunta este argumento es a la dificultad de ofrecer explicaciones de estas cuestiones que resulten compatibles con una imagen naturalista del mundo. Partiendo del reconocimiento de estas dificultades, lo que el argumento concluye es que para evitar tener que hacer frente a la necesidad de explicar la cuestiones mencionadas, la mejor alternativa es optar por renunciar al cognitivismo.

Desde la perspectiva de Smith, el principal problema de la estrategia expresivista es su incapacidad para acomodar nuestras intuiciones de sentido común con respecto a la naturaleza de los juicios morales. Pues más allá del hecho de que con base en nuestra experiencia haciendo uso de estos juicios simplemente parece correcto afirmar el cognitivismo, también debemos de tomar en cuenta que al suponer que los juicios morales no expresan más que estados de aprobación o desaprobación carentes de contenido proposicional, lejos de salir de problemas, nos enfrentamos con otras complicaciones tales como la de dotar de sentido a cualquier inferencia que tenga entre sus premisas a un juicio moral.[20]

Por su parte, la estrategia externista consiste en afirmar la verdad de las proposiciones (i) y (iii) y negar la de (ii). Así, lo que los defensores de esta estrategia sostienen es que la relación entre juicios morales y motivación para actuar es una relación contingente cuya satisfacción no depende de algún aspecto interno de los juicios morales sino más bien de factores externos a ellos. El argumento más relevante empleado para la defensa de esta perspectiva se centra en

[17] Estas reconstrucciones provienen de (Smith, 1994, pp. 154-156).
[18] Las formulaciones más influyentes del expresivismo son las de (Blackburn, 1984) y (Gibbard, 1993 y 2000).
[19] Al menos bajo una concepción tradicional y no-minimalista de los criterios de verdad.
[20] Esta cuestión concerniente a la dificultad para explicar cómo es que si los juicios morales expresan estados mentales no cognitivos pueden participar en inferencias es lo que en la literatura se denomina "problema Frege-Geach". Para exposiciones detalladas de dicho problema véanse (Schroeder, 2007) y (Woods, 2018).

el reconocimiento de la existencia de contraejemplos a la tesis internista según la cual del juzgar moralmente necesariamente se sigue motivación para actuar. La versión más influyente de este argumento es desarrollada originalmente por Brink (1989) y gira en torno a la figura de agentes denominados "amorales", quienes se caracterizan por realizar juicios morales sin por ello obtener algún tipo de motivación para actuar. Para el externismo, los casos de agentes amorales se articulan como contraejemplos a la tesis internista de que los juicios morales implican conceptualmente la motivación para actuar.

El problema que Smith encuentra con la estrategia externista fue descrito en la presentación de su argumento a la mejor explicación a favor del internismo moral. Se trata del hecho de que para dar cuenta de la motivación moral, el externista no tiene más opción que afirmar el fetichismo moral. Esto se debe a que, puesto que la tesis central del externismo es que la conexión entre juicio moral y motivación tiene un fundamento externo a los juicios morales, esta perspectiva no puede afirmar que la motivación que produce un juicio moral nunca se fundamenta directamente en el contenido del juicio

Finalmente, la estrategia anti-humeana consiste en afirmar la verdad de (i) y (ii) y rechazar la de (iii). La idea en torno a la cual opera esta estrategia es la de que las razones motivacionales, en su condición de estados psicológicos que permiten explicar la realización de acciones intencionales, tienen como único elemento constitutivo creencias. El argumento principal para defender esta idea apunta a la observación de que, cuando menos en algunas ocasiones, actuamos de una determinada manera simplemente porque creemos que esa es la forma más adecuada de actuar. Dada esta observación, desde la perspectiva anti-humeana se afirma que la tesis de que tanto creencias como deseos son condición necesaria para que haya motivación constituye un dogma. En particular, un dogma que reduce la acción intencional a una cuestión meramente instrumental.

Para Smith, el problema con la estrategia anti-humeana es su ignorancia con respecto a la distinción entre dos tipos de razones para la acción: razones motivaciones y razones normativas. Una vez que esta distinción es tomada en consideración, puede apreciarse la falsedad de los supuestos contraejemplos a la teoría humeana referidos por los anti-humeanos. Pues en realidad, dichos casos refieren a situaciones en las que se produce una convergencia entre las razones normativas y las razones motivacionales de un sujeto.

4. La solución racionalista de Smith

La propuesta de solución a PM de Smith se articula como una teoría de los juicios morales que afirma que estos juicios expresan creencias acerca de razones normativas para la acción. De acuerdo con la teoría, cuando un sujeto *S* juzga moralmente correcto hacer φ en las circunstancias *C*, *S* expresa una creencia acerca de la existencia de una razón normativa para hacer φ en *C*. La teoría de Smith es racionalista, en el sentido de que sostiene que el criterio normativo de la corrección moral es en última instancia un criterio normativo de racionalidad. Hacer lo moralmente correcto es actuar racionalmente y hacer lo moralmente incorrecto es actuar irracionalmente. Concebir de esta manera a los juicios morales permite, según Smith, ofrecer una solución conciliadora a PM. Es decir, permite afirmar que los juicios morales expresan estados mentales cognitivos y además sostienen un vínculo interno con la motivación para actuar, sin que ello implique negar la teoría humeana de la motivación.

Para el desarrollo de esta teoría racionalista de los juicios morales, Smith se dedica a analizar el concepto de razón normativa. El punto de partida de su análisis se ubica en la caracterización del tipo de explicación de la acción intencional que se puede dar en términos de razones normativas. A diferencia de las razones motivacionales, las cuales explican la acción intencional en términos del "complejo de estados psicológicos que producen una acción", las razones normativas explican la acción intencional especificando "el patrón de deliberación que produjo o pudo haber producido una acción" (Smith, 1994, p. 131).[21] Dada su fundamentación en la deliberación, Smith (1994, p. 131) caracteriza a la explicación de la acción intencional en términos de razones normativas como una explicación desde una perspectiva deliberativa. Al deliberar, reflexionamos acerca de hechos que podrían justificar racionalmente una acción. Y son las proposiciones que capturan esos hechos, capaces de justificar acciones desde la racionalidad, aquello a lo cual reconocemos como razones normativas. De esta manera, una razón normativa explica la acción intencional de un sujeto en la medida en que, desde la perspectiva de ese sujeto, dicha razón justifica su acción. Pues al menos en algunos casos, cuando a través de la deliberación determinamos qué acción tendríamos justificación de llevar a cabo, esa determinación hace una diferencia en qué es lo que de hecho terminamos haciendo.

[21] De acuerdo con Smith (1994, p.132), es importante reconocer que no siempre que actuamos intencionalmente atravesamos un proceso deliberativo que produce la acción relevante. En ocasiones, con el objetivo de justificar *ex post facto* una acción, simplemente suponemos cuál podría haber sido ese proceso deliberativo.

Pero si la deliberación puede tener un resultado práctico, en el sentido de que puede resultar en la realización de acciones, entonces habrá que especificar qué tipo de actitud proposicional produce la deliberación. Pues solo así se podrá saber si la explicación de la acción intencional en términos de razones normativas es compatible con la explicación de la acción intencional en términos de razones motivacionales, entendidas estas últimas en apego a lo estipulado por la teoría humeana de la motivación (THM). En concreto, solo especificando qué tipo de actitud proposicional produce la deliberación se podrá saber si al explicar la acción intencional en términos de razones normativas no estamos contradiciendo la tesis de que poseer una creencia y un deseo es condición necesaria y suficiente para tener motivación. De acuerdo con Smith (1994, p. 132), cuando un sujeto reconoce, a través de la deliberación, que una acción en particular cuenta con la justificación de una razón normativa, la actitud que dicho sujeto posee puede denominarse "valorar". Pensar que hay una razón normativa para hacer φ consiste en *valorar* hacer φ. Sin embargo, por sí sola, la identificación de la actitud resultante de la deliberación como *valorar* no permite indicar cómo se relaciona esta actitud con las razones motivacionales.

Con el objetivo de esclarecer la naturaleza del *valorar*, Smith procede a explorar la posibilidad de entender esta actitud en términos de otra actitud proposicional. Primero, siguiendo propuestas en la literatura como las de Davidson (1978), Gauthier (1986) y Lewis (1989), contempla la posibilidad de que *valorar* consista en *desear*. Sin embargo, partiendo del reconocimiento de la existencia tanto de casos en los que se desea lo que no se valora, como de casos en los que no se desea lo que se valora, descarta semejante concepción.[22] Es así como Smith pasa a abordar la posibilidad de que *valorar* consista en *creer*. A diferencia de lo que sucedía con los deseos, la concepción de valorar como creer no parece enfrentar contraejemplos que cuestionen su plausibilidad. Pero eso no significa que no enfrente dificultades. En particular, siguiendo a Lewis (1989, p. 116), Smith considera que la viabilidad de una concepción semejante dependerá de la posibilidad de mostrar que puede haber una conexión conceptual entre *creer* y *desear*. Pues solo si una conexión de este tipo es el caso, podrá afirmarse que la explicación de la acción intencional en términos de razones normativas es compatible con THM.

[22] En lo correspondiente a los casos en los que se desea lo que no se valora, Smith (1994, p. 133) cita el caso de un cleptómano descrito por Ayer (1954, p. 20), el de un adicto a la heroína descrito por Frankfurt (1971, p. 87) y un par de casos planteados por Watson (1975, p. 101). Y en lo concerniente a los casos en los que no se desea lo que se valora, Smith (1994, p. 135) se remite a casos descritos por Stocker (1979, p. 744), en los que la desconexión entre las actitudes relevantes se produce como como consecuencia de factores emocionales o físicos tales como la depresión, el cansancio extremo, etc.

La estrategia de Smith para mostrar que efectivamente puede haber una conexión conceptual entre creer y desear se articula como una explicación del siguiente principio:

> **C2**: "Si un sujeto cree que tiene una razón normativa para hacer *φ*, entonces racionalmente debería desear hacer *φ*" (Smith, 1994, p. 148).[23]

De acuerdo con Smith, para entender C2 hay que retomar el análisis del concepto de razón normativa. En particular, hay que llamar la atención sobre una trivialidad acerca de dicho concepto que Korsgaard (1986) ya ha reconocido con anterioridad. La trivialidad en cuestión es que, aquello para lo cual creemos que tenemos una razón normativa, es lo que pensamos que desearíamos hacer si fuéramos plenamente racionales. Al creer que tenemos una razón normativa para hacer φ, creemos que, desde la perspectiva de una versión plenamente de nosotros mismos, sería deseable que dadas nuestras circunstancias actuales deseáramos hacer φ. Lo que subyace a esta trivialidad es simplemente el reconocimiento del hecho de que, en la medida en que las creencias acerca de razones normativas son resultado del despliegue de ejercicios deliberativos, formar una creencia verdadera de ese tipo implicaría hacer un uso perfecto de la deliberación. Es precisamente para clarificar qué es lo que supondría semejante uso perfecto de la deliberación que Smith, siguiendo a Williams (1980), se remite a la figura de un agente plenamente racional. De acuerdo con Williams (1980), un agente plenamente racional tendría que ser capaz de satisfacer las siguientes tres condiciones:

(1) El agente no debería de poseer creencias falsas.
(2) El agente debería de poseer todas las creencias verdaderas relevantes.
(3) El agente debería de deliberar correctamente.

Para Smith, Williams está en lo cierto al identificar esas tres condiciones como aquellas que un agente plenamente racional debería de cumplir. Sin embargo, más allá de la especificación de las condiciones, aunque en lo concerniente a (1) y (2) su acuerdo con Williams es total, en lo particular de la caracterización de (3), Smith toma cierta distancia.

Tanto para Williams como para Smith, la deliberación es una actividad que consiste en crear y destruir deseos. Ambos coinciden en que esta actividad involucra una pluralidad de operaciones. Sin embargo, el punto de desacuerdo se encuentra en la identificación del tipo de

[23] La decisión de nombrar "C2" a este principio responde al interés por preservar los términos en los que Smith plantea la cuestión.

operación que más peso debería de tener en la deliberación de un agente plenamente racional. Desde la perspectiva de Williams, la operación relevante es el ejercicio de la imaginación. Si un sujeto imagina adecuadamente todos los desarrollos posibles que podría tener una situación, lo que cabrá esperar es que lo que termine deseando y, por tanto, haciendo, de hecho sea lo más racional que pudo haber deseado. Para Smith (1994, p.158), aunque ciertamente el ejercicio de la imaginación debería desempeñar un rol importante, no se trata de la operación que debería de regir la deliberación de un agente plenamente racional. Más bien, la operación protagónica en un contexto idealizado semejante debería de ser "el intento de determinar si nuestros deseos están *sistemáticamente justificados*" (Smith, 1994, pp.158-159).[24] Al determinar qué desear, un agente plenamente racional debería de priorizar la búsqueda de aquellos deseos que al integrarse dentro del conjunto de sus deseos, doten de mayor coherencia y unidad al conjunto en su totalidad. Semejante búsqueda de coherencia y unidad es la operación más racional porque es la que permite construir un conjunto sistemáticamente justificado de deseos. Es decir, un conjunto en el que la satisfacción de deseos derivados más específicos permite satisfacer a deseos no-derivados más generales.

Es con base en la caracterización anterior de la deliberación práctica que Smith explica por qué un principio como C2 es verdadero y por tanto, por qué se puede afirmar que efectivamente existe una conexión conceptual entre el *creer* propio del *valorar* y *desear*. Si un sujeto cree que tiene una razón normativa para hacer φ, racionalmente debería desear hacer φ porque eso es lo que le permitirá dotar de mayor coherencia y unidad al conjunto de sus deseos. Hay, pues, una demanda normativa desde la racionalidad práctica que hace que de su *creer* deba de seguirse un *desear*.

Con base en la exposición anterior de la teoría racionalista de los juicios morales de Smith, podemos proceder a especificar por qué dicha propuesta puede tomarse como una solución conciliadora a PM. Para ello, indicaré de qué manera la teoría explica cada una de las proposiciones que conforman el problema.

La teoría de Smith puede dar cuenta de la proposición (i) de PM en tanto se articula como una forma de cognitivismo moral. En concreto, la teoría es cognitivista porque afirma que los juicios morales expresan un cierto tipo de creencias, específicamente, creencias acerca de razones normativas para la acción. Puesto que las razones normativas son

[24] De acuerdo con Smith, esta imagen de la deliberación práctica está inspirada en la idea de "equilibrio reflexivo" entre creencias evaluativas descrita por Rawls (1951).

proposiciones que capturan hechos objetivos respecto a la justificación de acciones, en última instancia, la teoría de Smith constituye una forma de cognitivismo moral objetivista.

Por otro lado, la teoría de Smith puede dar cuenta de la proposición (ii) de PM porque se articula como una forma de internismo moral. La teoría es internista porque contempla que, en lo particular de los agentes racionales en términos prácticos, la emisión de un juicio moral implica motivación para actuar. A través del análisis del concepto de razón normativa, Smith muestra que las creencias expresadas por juicios morales son creencias acerca de lo que la versión plenamente racional de uno mismo desearía que uno hiciera. Posteriormente, al mostrar que, dada la naturaleza de la correcta deliberación, uno debería racionalmente desear como cree que su versión plenamente racional desearía que lo hiciera, Smith clarifica el sentido en el cual, si uno es racional, entonces, al emitir un juicio moral tendrá motivación para actuar.

Finalmente, la teoría de Smith puede dar cuenta de la proposición (iii) de PM en tanto es compatible con la afirmación de las dos tesis propias de THM. Al mostrar que para generar motivación, las creencias expresadas por juicios morales tienen que sostener una conexión conceptual con la formación de deseos, la teoría confirma su compatibilidad con la tesis débil de THM. Y puesto que en su explicación de la motivación moral todo el trabajo motivacional es desempeñado únicamente por una creencia y un deseo, la teoría es compatible con la verdad de la tesis fuerte de THM.

5. Conclusiones

Este capítulo ha estado dedicado a cumplir dos objetivos fundamentales. El primero de esos objetivos ha sido ofrecer una reconstrucción del trilema que Smith (1994) denomina el "problema moral". Para ello, he especificado qué es exactamente lo que subyace a cada de las proposiciones que participan del problema. En lo concerniente a la proposición (i), la exposición estuvo centrada en la descripción del cognitivismo moral. Posteriormente, en lo relativo a la proposición (ii), la atención pasó a la caracterización del internismo moral. Y en el caso de la proposición (iii), me dediqué a identificar los compromisos inherentes a la teoría humeana de la motivación.

El segundo objetivo que este capítulo cumplió fue el de revisar diferentes tipos de soluciones al problema moral. Primero, siguiendo el análisis elaborado por Smith acerca de las soluciones existentes en la literatura, identifiqué tres soluciones parciales al problema: la

expresivista, la externista y la anti-humeana. Y después, tras subrayar los problemas que afectan a cada una de esas soluciones, pasé a reconstruir la propuesta de solución conciliadora desarrollada por Smith. Dicha propuesta se articula como una teoría racionalista de los juicios morales. La tesis central que defiende es que al emitir un juicio moral, un sujeto expresa su posesión de una creencia acerca de la existencia de una razón normativa para la acción.

CAPÍTULO II

Objeciones a la solución racionalista de Smith al problema moral

1. Introducción

El capítulo I de este trabajo estuvo dedicado a reconstruir tanto el problema moral (PM) formulado por Smith (1994) como la solución que el mismo Smith plantea a dicho problema. De acuerdo con lo indicado a lo largo del capítulo, PM es un problema que concierne a la inconsistencia entre tres enfoques teóricos aparentemente plausibles: el cognitivismo moral, el internismo moral y la teoría humeana de la motivación. Por su parte, la solución propuesta por Smith se articula como una teoría racionalista de los juicios morales, según la cual estos juicios expresan creencias acerca de razones normativas para la acción.

En este segundo capítulo, pasaré a centrar la atención sobre objeciones a la teoría de Smith. El capítulo está estructurado de la siguiente manera. En la sección 2 me dedicaré a clarificar qué es exactamente lo que se entiende por juzgar moralmente. Posteriormente, en la sección 3, identificaré con precisión los diferentes elementos que dan forma a la teoría racionalista de Smith. Y finalmente, en la sección 4, mostraré cómo es que dados los elementos que la componen, la teoría de Smith enfrenta tres objeciones diferentes. Una primera objeción atribuye a la teoría un problema de hiper-intelectualización. Después, una segunda objeción refiere a una limitación explicativa inherente a los términos en los que la teoría concibe la normatividad de los juicios morales. Y por último, una tercera objeción apunta a la incapacidad de la teoría para ofrecer una explicación del fenómeno de la motivación moral que sea compatible con el internismo moral.

2. Un caso claro de juicio moral

A continuación, procederé a describir un caso que permita indicar con precisión en qué consiste emitir un juicio moral. La descripción de este caso no sólo servirá para delimitar el fenómeno que protagoniza PM, sino que además funcionará como ejemplo para ilustrar cómo opera la explicación de este fenómeno desde la teoría racionalista de Smith.

Miguel es una persona racional que en el curso de su vida cotidiana procura ser amable y respetar a los demás. Hace unos días, mientras desayunaba y veía las noticias, se enteró de

que la empresa dueña de una marca de ropa muy popular recurre a prácticas de explotación laboral para la elaboración de sus prendas. Durante el noticiero, una periodista reputada por la fiabilidad de sus investigaciones, presentó un reportaje dedicado a exponer evidencia que parece mostrar que la mencionada empresa cuenta con fábricas clandestinas en las que personas trabajan turnos de 20 horas al día, sin tiempos de descanso y a cambio de remuneraciones económicas muy por debajo de lo establecido por la ley. Al enterarse de la noticia, Miguel ha experimentado una gran indignación. Desde su punto de vista, la explotación laboral y, por tanto, la conducta de la empresa, son moralmente incorrectas. Considera que la explotación constituye una práctica cimentada en una relación injusta, en la que una persona o grupo obtiene beneficios aprovechándose de la situación de vulnerabilidad de las personas que son explotadas.

La emisión de semejante juicio de incorrección moral con respecto al hecho referido en las noticias ha tenido un cierto efecto en la motivación para actuar de Miguel. En concreto, lo ha motivado a no comprar prendas elaboradas por la marca de ropa de la que se hablaba en el reportaje. Tal vez, si Miguel estuviera menos ocupado o fuera más enérgico en la defensa de sus convicciones morales, no solo tendría motivación para ya no comprar prendas de la mencionada marca de ropa, sino también para organizar una manifestación o para preparar una demanda. Pero independientemente de la radicalidad específica de aquella acción para la cual ha adquirido motivación, el hecho relevante es que Miguel se siente motivado para actuar de una manera que se ajusta al contenido de su juicio moral según el cual la explotación realizada por la empresa en cuestión es incorrecta. Si por el contrario, juzgara la explotación laboral como una práctica moralmente correcta, entonces, seguramente las acciones para las cuales estaría motivado serían otras. En concreto, serían acciones ajustadas a ese juicio moral alternativo. Y si en cambio, Miguel no tuviera un juicio moral con respecto a la explotación laboral, entonces el reportaje que vio en el noticiero no tendría ningún tipo de efecto sobre su motivación para actuar. Pues por sí sólo, el conocimiento de un hecho empírico tal como el de que una cierta empresa explota a sus trabajadores no genera motivación para actuar de una forma en particular.

3. Los elementos de la teoría racionalista de Smith

Recordando la descripción realizada en la sección 4 del capítulo anterior, la teoría racionalista de los juicios morales elaborada por Smith puede esquematizarse en una tabla como la que presento a continuación:

Elemento	Descripción
(A) Conjunto de estados mentales	Si un sujeto juzga moralmente, entonces dicho sujeto posee dos estados mentales diferentes. Los estados en cuestión son: **A1.** Una creencia acerca de lo que sería deseable que uno hiciera en sus circunstancias actuales desde la perspectiva de una versión plenamente racional de uno mismo. **A2.** Un deseo que corresponde con lo que se cree deseable en *A1*.
(B) Tesis de identidad	El concepto de lo deseable desde la perspectiva de la plena racionalidad es analíticamente idéntico al concepto de corrección moral.
(C) Imagen de la plena racionalidad práctica	La plena racionalidad práctica consiste en deliberar correctamente, entendida como la capacidad El concepto idealizado de plena racionalidad se define como la conjunción de tres capacidades: **C1.** La posesión de todas las creencias verdaderas relevantes dadas ciertas circunstancias. **C2.** La no posesión de creencias falsas. **C3.** La correcta deliberación, entendida como la habilidad de mantener la coherencia y la unidad dentro del conjunto de los estados mentales de uno mismo.

4. Objeciones a la teoría racionalista

A lo largo de esta sección presentaré tres objeciones a la teoría racionalista de los juicios morales que Smith plantea como solución a PM. Cada una de las objeciones ataca un aspecto diferente de la teoría. En la sección 4.1, desarrollaré una objeción que, partiendo del reconocimiento del tipo de estados mentales que de acuerdo con la teoría de Smith expresamos al emitir un juicio moral, procede a afirmar que dicha teoría enfrenta un problema de hiper-intelectualización. Posteriormente, en la sección 4.2, formularé una objeción que apunta a un problema inherente a los términos en los que la teoría de Smith conceptualiza el tipo de justificación que los juicios morales pueden dar a una acción. Y finalmente, en la sección 4.3, presentaré una objeción que pone en duda la capacidad de la teoría de Smith para explicar, en apego al internismo moral, el fenómeno de la motivación moral. En última instancia, cada una de estas objeciones permite extraer una lección importante con respecto a cómo debería de ser una teoría adecuada de los juicios morales.

4.1 Objeción de hiper-intelectualización

La primera objeción a la teoría racionalista de Smith que abordaré tiene su origen en conversaciones sostenidas con mi tutor en el contexto de la elaboración de esta tesis. El nombre que le he dado es "objeción de hiper-intelectualización", pues lo que afirma es que hay un problema de hiper-intelectualización inherente a la imagen del juzgar moralmente que se desprende de la teoría de Smith. De acuerdo con esta teoría, al emitir un juicio moral, un sujeto expresa su posesión de un estado mental como el que en el esquema elaborado en la sección 3 he identificado bajo la etiqueta de (A1). Es decir, una creencia acerca de lo que la versión plenamente racional de sí mismo desearía que hiciera en su versión actual.

La objeción de hiper-intelectualización puede desarrollarse a través de dos vertientes diferentes. Una primera vertiente apunta a la incompatibilidad entre la complejidad de un estado mental como el descrito en (A1) y el carácter automático y rápido de los juicios morales. Esta incompatibilidad resulta problemática en tanto parece mostrar que la teoría de Smith ofrece una imagen falsa del tipo de procesamiento cognitivo requerido para juzgar moralmente. Por otra parte, una segunda vertiente de la objeción se centra en el hecho de que al suponer que la posesión de un estado como (A1) es una condición necesaria para emitir un juicio moral, la teoría de Smith se compromete con la afirmación implausible de que si un

sujeto juzga moralmente, entonces posee conceptos de un alto grado de complejidad. En este caso, el aspecto problemático es consecuencia de la existencia de casos en los que parece razonable afirmar que, sujetos que no poseen conceptos de un alto grado de complejidad, pueden emitir juicios morales. Independientemente de la vertiente a través de la cual se desarrolle, lo que esta objeción muestra es que la teoría racionalista de Smith hiper-intelectualiza las condiciones cognitivas que habilitan la posibilidad de llevar a cabo un juicio moral. A continuación, procederé a desarrollar la objeción en sus dos vertientes.

Para ilustrar la primera vertiente de la objeción de hiper-intelectualización, retomaré el caso de Miguel descrito en la sección 2 de este capítulo. Desde la teoría racionalista de Smith, un caso semejante se explicaría en los siguientes términos. Al escuchar la noticia de que la empresa explota a sus trabajadores, Miguel emitió un juicio de incorrección moral y, con ello, expresó su posesión de una cierta creencia. En concreto, expresó su posesión de una creencia de acuerdo con la cual, la versión plenamente racional de sí mismo desearía que él, en su versión actual, evite que la empresa siga explotando trabajadores. Es el contenido de esa creencia acerca de lo que una versión plenamente racional de sí mismo desearía que hiciera lo que, según la teoría de Smith, permite explicar por qué, el juicio moral de Miguel, lo ha llevado a adquirir motivación para no consumir productos elaborados por la empresa en cuestión. Como mencioné previamente, el aspecto problemático sobre el cual llama la atención esta primera vertiente de la objeción es la existencia de una cierta incompatibilidad entre la imagen recién descrita del juzgar moralmente y la naturaleza de la experiencia que de hecho tenemos cuando emitimos un juicio moral. Es decir, una incompatibilidad entre la teoría de Smith y la fenomenología de los juicios morales.

El rasgo de la fenomenología de los juicios morales con el cual no parece ser compatible la teoría de Smith es el carácter rápido y automático propio de los juicios morales. En general, cuando emitimos un juicio moral, la experiencia que tenemos es más similar a la de percibir un objeto en nuestro entorno que a la de reflexionar detenidamente acerca de qué deberíamos hacer. Se trata, pues, de una experiencia breve, que parece involucrar poco o nulo control sobre nuestros pensamientos. Para ilustrar esta caracterización de la experiencia de emitir un juicio moral podemos remitirnos nuevamente al caso de Miguel. Al pensar en el momento en el que Miguel emite su juicio de incorrección moral con respecto a la acción de la empresa de explotar a sus trabajadores, lo más natural será suponer, sobre la base de nuestra propia experiencia emitiendo juicios morales, que la experiencia de Miguel no ha abarcado más que unos cuantos segundos. Emitió su juicio tan pronto escuchó la noticia. Por

el contrario, al pensar en el tipo de experiencia que un sujeto tendría al preguntarse por qué es lo que una versión plenamente racional de sí mismo desearía dadas sus circunstancias actuales, lo más probable es que supongamos que la experiencia en cuestión será larga. Inclusive, puesto que lo que se está preguntando es algo que por su propia naturaleza únicamente podría resolverse de manera definitiva desde una perspectiva idealizada, sería razonable suponer que la experiencia en cuestión se extendería indefinidamente en el tiempo. Si bien este contraste fenomenológico puede parecer trivial, en tanto el fenómeno respecto del cual se está teorizando consiste, en parte, en un tipo de experiencia, la dimensión fenomenológica de la cuestión no debería de pasarse por alto tan fácilmente. En cualquier caso, la relevancia de esta primera vertiente de la objeción se debe a que permite llamar la atención sobre el hecho de que la teoría de Smith implica una imagen implausible del tipo de procesamiento cognitivo que posibilita la emisión de juicios morales.

Ahora bien, como indiqué con anterioridad, la segunda vertiente de la objeción de hiper-intelectualización apunta a que la teoría racionalista de Smith no puede acomodar el hecho de que sujetos que no poseen conceptos de un alto grado de complejidad, de hecho pueden emitir juicios morales. Para ilustrar esta segunda vertiente, procederé a plantear un nuevo ejemplo. Pensemos en el caso de María, la hija de cinco años de Miguel, quien mientras su papá veía el noticiero se encontraba en la escuela. Durante el recreo, María se percató de que algunos de sus compañeros molestaban a un niño más pequeño por su baja estatura. Al ver la escena, María sintió que lo que estaba presenciando no estaba bien. De hecho, experimentó muchas ganas de ir con su maestra para hablarle de lo que estaba sucediendo, tantas ganas que finalmente salió corriendo en búsqueda de su maestra.

Frente a un caso como el de María, parecería razonable afirmar que lo que ha llevado a María a buscar hablar con su maestra es la motivación proveniente de su juzgar la incorrección moral de la conducta de sus compañeros. Sin embargo, si como supone la teoría de Smith, para haber emitido un juicio moral, María tuvo que haber desplegado un ejercicio deliberativo que involucró el manejo de conceptos de un alto grado de complejidad, entonces, la posibilidad de atribuir a María un juicio moral quedará socavada. Algunos de los conceptos complejos que, según la teoría de Smith, María tendría que haber podido manejar son: “racionalidad”, “plena racionalidad”, “agente plenamente racional” y “agente plenamente racional que es una versión de mí misma”. Partiendo del reconocimiento de estos conceptos, podemos percatarnos de que, dado el nivel de abstracción que suponen, resulta implausible que una niña de cinco años como María de hecho los posea.

Sobre la base de la exposición anterior de las dos vertientes de la objeción de hiper-intelectualización, podemos llegar a la siguiente lección con respecto a lo que cabría esperar de una buena teoría de los juicios morales:

> ***Plausibilidad psicológica:*** una buena teoría de los juicios morales debería (i) acomodar la aparente rapidez y automaticidad de muchos juicios morales y (ii) no exigir la posesión de conceptos que no estén disponibles para ciertos sujetos capaces de formar juicios morales.

4.2. Objeción de justificación contingente

La segunda objeción a la teoría de los juicios morales de Smith que presentaré tiene su origen en ideas desarrolladas originalmente por Dancy (2004, pp. 53-70). En particular, encuentra su punto de partida en la distinción planteada por Dancy entre razones contributivas y razones globales.[25] A esta objeción la denomino "objeción de justificación contingente", pues se centra en la idea de que la teoría de Smith es equivocada en tanto no puede acomodar el hecho de que la justificación de la acción que un juicio moral puede ofrecer es contingente. Dicha contingencia se hace patente en el hecho de que, como consecuencia de la interacción con otros juicios morales, la justificación de la que dota un juicio moral puede fortalecerse, debilitarse o incluso cortarse. El motivo por el cual la teoría racionalista de Smith no puede acomodar esta contingencia es porque según establece, los juicios morales se articulan como veredictos finales desde la perspectiva de la plena racionalidad. Es decir, se trata de una consecuencia inherente a la tesis de identidad plasmada en el elemento (B) del esquema de la teoría de Smith presentado en la sección 3 de este capítulo.

La naturaleza contingente del tipo de justificación de la acción que ofrecen los juicios morales es algo que podemos constatar en nuestra experiencia cotidiana emitiendo este tipo de juicios. Para ilustrar la cuestión, me remitiré nuevamente al caso de Miguel descrito en la sección 2 del presente capítulo. En concreto, propondré tres variaciones a ese caso que permiten apreciar tres sentidos diferentes en los cuales la justificación de la acción que ofrece un juicio moral puede cambiar.

[25] La noción de "razones globales" se utiliza como traducción de la noción original en inglés empleada por Dancy (2004) de "all-things-considered reasons".

La primera variación al caso de Miguel que presentaré, tiene el objetivo de ilustrar un tipo de cambio de la justificación que da un juicio moral que consiste en el *fortalecimiento* de dicha justificación. Supongamos que poco después de enterarse de la noticia de explotación laboral, Miguel contactó con una nueva noticia en la que se presentó evidencia acerca de un enorme daño ambiental generado por las fábricas de la misma empresa que protagonizaba la noticia de explotación. Puesto que es una persona preocupada por el medio ambiente y el futuro del planeta, al enterarse de este hecho, Miguel emitió un nuevo juicio de incorrección moral. Al igual que su juicio acerca de la explotación laboral, este segundo juicio de incorrección moral justifica una acción como la de nunca consumir productos elaborados por la marca en cuestión. Pues es ese tipo de acción la que se ajusta al nuevo juicio moral emitido por Miguel. Pero entonces, dada la justificación derivada de su segundo juicio, la justificación que, sobre la base de su juzgar incorrecta la explotación laboral, Miguel tenía para no comprar productos de la marca, ha cambiado. En concreto, puesto que ahora cuenta con el apoyo de dos razones diferentes, dicha justificación se ha fortalecido.

Semejante fortalecimiento de la justificación que Miguel tiene para no comprar productos elaborados por la empresa que explota a sus trabajadores y contamina el ambiente es algo que la teoría de los juicios morales de Smith no puede acomodar. Pues de acuerdo con esa teoría, en tanto la justificación que ofrece un juicio moral refleja un veredicto final desde la perspectiva de la plena racionalidad, esa justificación no es susceptible al cambio.

La segunda variación al caso de Miguel que presentaré, tiene el objetivo de ilustrar un tipo de cambio de la justificación que da un juicio moral que consiste en el *debilitamiento* de dicha justificación. Imaginemos que, previo a contactar con la noticia de explotación laboral, Miguel tuviera conocimiento acerca del hecho de que la empresa de la que se hablaba en las noticias lleva años haciendo donaciones sustanciales para el fomento de investigación científica que ha permitido desarrollar medicamentos para combatir el cáncer. Inclusive, a través del testimonio de un buen amigo suyo que trabaja como médico, se ha enterado de que la investigación financiada por la empresa ha permitido elaborar medicamentos que han salvado la vida a miles de personas. Desde la perspectiva de Miguel, esta práctica de donación realizada por la empresa es moralmente correcta. Sin embargo, ello no cambia el hecho de que juzga a la explotación laboral como una práctica moralmente incorrecta. Por lo tanto, cuando se ha enterado de que la empresa en cuestión explota a sus trabajadores, ha emitido un juicio de incorrección moral. Como resultado de este choque entre juicios

morales, la justificación que, sobre la base de su juicio inicial de corrección moral, Miguel tenía para llevar a cabo ciertas acciones, ha cambiado. En un primer momento, cuando solo estaba enterado de la práctica de donación, Miguel tenía justificación para comprar productos elaborados por la empresa. Pero ahora, tras informarse de la práctica de explotación y dado su juicio moral de incorrección, esa justificación para comprar productos de la empresa ha disminuido. Pues aunque Miguel sigue pensando que su juicio de corrección moral justifica comprar productos de la empresa, reconoce que al actuar de esa manera está yendo en contra de las acciones que justifica su juzgar moralmente incorrecta la explotación laboral.

Desde la teoría racionalista de Smith, un caso como el recién descrito no podrá explicarse en términos de un cambio en la justificación. Pues nuevamente, según lo establecido por dicho enfoque, en tanto el juicio de corrección moral de la práctica de donación, como el juicio de incorrección de la práctica de explotación, reflejan un veredicto final desde la plena racionalidad, la justificación que ofrecen no es susceptible a cambiar.

Finalmente, la tercera variación al caso de Miguel que presentaré, tiene el objetivo de ejemplificar un tipo de cambio de la justificación que da un juicio moral que consiste en el *corte* de dicha justificación. Supongamos que, tiempo después de escuchar la noticia de explotación laboral, Miguel se enterase, a través de una investigación publicada por otro periodista fiable, de que dada la precariedad de su contexto, los trabajadores empleados por la empresa en cuestión no cuentan con ninguna otra opción de trabajo. De hecho, la gravedad de la situación es tal que, en caso de que los trabajadores pierdan el trabajo en el que son explotados, lo más probable es que queden en la ruina total y se queden sin acceso a los bienes más elementales para subsistir. Para Miguel, promover, aunque sea indirectamente, que una persona pierda acceso a los bienes necesarios para sobrevivir es, en definitiva, moralmente incorrecto. Por lo tanto, ha emitido un nuevo juicio moral de acuerdo con el cual, no consumir productos de la empresa referida en las noticias y con ello promover que los trabajadores pierdan su trabajo, sería moralmente incorrecto. Como consecuencia de este segundo juicio moral, la justificación que inicialmente tenía Miguel para no comprar productos de la empresa explotadora ha cambiado. En particular, esa justificación original ha sido cortada, en el sentido de que ha quedado neutralizada por el nuevo juicio moral emitido por Miguel.

El fenómeno de corte de la justificación ofrecida por un juicio moral descrito a través de esta última variación al caso de Miguel, es algo respecto de lo cual la teoría de los juicios morales de Smith no podría dar cuenta. Pues una vez más, al suponer que la normatividad de

los juicios morales se articula bajo la forma de veredictos finales desde la perspectiva de la plena racionalidad, la teoría se compromete con la afirmación de que la justificación que ofrece un juicio moral no es susceptible a cambiar.

Partiendo de las variaciones al caso de Miguel recién descritas, queda claro cuál es el problema planteado por la "objeción de motivación contingente". Al conceptualizar a los juicios morales en términos de creencias sobre lo que un agente plenamente racional desearía, Smith atribuye a estos juicios la condición de razones globales. Es decir, los concibe como veredictos finales, en el sentido de que son el resultado de haber tomado en cuenta todas las consideraciones potencialmente relevantes para determinar qué acción sería deseable dadas ciertas circunstancias. Pero si esto es así, entonces, la teoría de Smith no puede acomodar el hecho de que la justificación de la acción que produce un juicio moral puede cambiar. En particular, no puede acomodar el hecho de que esa justificación puede fortalecerse, disminuirse o cortarse. Por lo tanto, la objeción parece llevar a una disyuntiva: o bien el cambio en la justificación que dan los juicios morales es un fenómeno real, o bien la teoría racionalista de los juicios morales de Smith es verdadera. Dada esta disyuntiva, puesto que el cambio en la justificación que dan los juicios morales parece confirmarse sistemáticamente en nuestra experiencia, lo más razonable será concluir que la teoría de Smith no es verdadera. Pues dicha teoría no ofrece ningún rendimiento explicativo capaz de compensar la negación de lo que en el marco de la experiencia cotidiana resulta evidente. Así, la objeción de justificación contingente también arroja una lección importante acerca de lo que hay que esperar de una teoría adecuada de los juicios morales. La lección en cuestión es la siguiente:

> ***Cambio en la justificación:*** una buena teoría de los juicios morales debería de poder acomodar el hecho de que la justificación que un juicio moral puede dar a una acción es susceptible de fortalecerse, disminuirse o cortarse como consecuencia de la interacción con otro juicio moral.

4.3. Objeción de modo de presentación inadecuado

La tercera objeción a la teoría racionalista de Smith que abordaré tiene su origen en consideraciones planteadas inicialmente por Sayre-McCord (1997) y que posteriormente he

desarrollado en intercambios con mi tutor. Ya que aquello en lo cual se centra es en una cierta inadecuación del modo de presentación propio del contenido de las creencias que según la teoría de Smith expresan los juicios morales, el nombre que le he dado es el de "objeción de modo de presentación inadecuado". El concepto de modo de presentación tiene su origen en la distinción entre sentido y referencia trazada originalmente por Frege (1892). A grandes rasgos, aquello sobre lo cual llama la atención este concepto es el hecho de que la manera específica en que un sujeto se presenta a sí mismo un contenido o proposición afecta el impacto que ese contenido o proposición tiene sobre el sujeto. Un ejemplo utilizado de manera recurrente en la literatura para ilustrar la cuestión es protagonizado por Luisa Lane, quien se encuentra profundamente enamorada de Superman. El enamoramiento de Luisa se refleja en el hecho de que cada vez que ve a Superman, experimenta una gran motivación para acercarse a conversar con él. Sin embargo, aunque Luisa no lo sabe, en muchas ocasiones, pese a encontrarse frente a Superman, no experimenta la motivación para acercarse propia de su estar enamorada. Esto se debe a que en esas ocasiones, el modo de presentación con el que Luisa identifica a Superman no es el de "Superman", sino el de "Clark Kent". Así, aunque el sujeto referido en todos los casos es el mismo, en la medida en que cambia el modo de presentación a través del cual Luisa identifica a ese sujeto, su reacción también se ve modificada.

En última instancia, a lo que apunta esta objeción es a que, dado el modo de presentación del contenido de las creencias que según la teoría de Smith expresan los juicios morales, el tipo de explicación que dicha teoría puede ofrecer de la motivación que produce un juicio moral es incompatible con la explicación que el internismo moral ofrece de ese mismo fenómeno. La razón por la cual esto es así es porque el modo de presentación en cuestión implica una desconexión entre el sujeto que emite un juicio moral y el fundamento de la motivación que genera ese juicio moral.[26] El aspecto problemático de esta incompatibilidad se debe al hecho de que como indiqué en el capítulo I de este trabajo, una solución conciliadora a PM como la que Smith pretende llevar a cabo tendría que poder afirmar las tres proposiciones que conforman el problema. Y puesto que el internismo moral corresponde con la proposición (ii) de PM, si la teoría de Smith no es compatible con su verdad, entonces tampoco podrá funcionar como solución conciliadora a PM.

Para desarrollar esta objeción hay que comenzar por recordar cuáles son los términos de la

[26] Como señalé en el capítulo I, la tesis principal del internismo es que existe una conexión entre juicio moral y motivación para actuar que tiene un fundamento interno a los juicios morales.

explicación que la teoría de Smith ofrece acerca de la capacidad de los juicios morales para motivar la acción. Retomando el elemento (A) en el esquema de la teoría de la teoría de Smith presentado en la sección 2 del presente capítulo, podemos apreciar que dicha teoría concibe a la motivación que genera un juicio moral como el resultado de la combinación de dos estados mentales diferentes. Por un lado, una creencia que corresponde con el tipo de creencias que según la teoría son expresadas por juicios morales. Es decir, una creencia como la que se describe en (A1), acerca de lo que sería deseable que uno hiciera en sus circunstancias actuales desde la perspectiva de una versión plenamente racional de sí mismo. Y por otro lado, un deseo, como el que se describe en (A2), que corresponda con lo que en (A1) se cree que es deseable. De acuerdo con Smith, en la medida en que un sujeto racional en términos prácticos debería desear hacer lo que cree que es deseable, al menos en lo particular de sujetos racionales, la creencia en (A1) sostiene una relación interna con el deseo en (A2). Así, si un sujeto racional tiene una creencia tipo (A1), también tendrá un deseo tipo (A2). Sin embargo, una vez que nos percatamos de una particularidad del contenido de (A1), la viabilidad de afirmar semejante relación interna se ve seriamente amenazada. La particularidad en cuestión es el modo de presentación de sí mismo a través del cual el sujeto identifica lo deseable. En concreto, dicho modo de presentación no es el de él, en su versión actual, sino el de una versión idealizada él. Como consecuencia del hecho de que lo deseable en la creencia (A1) se especifique bajo ese modo de presentación, se produce una distancia entre el sujeto en su versión actual y lo que cree deseable que impide que se pueda afirmar, como Smith pretende, que hay una relación interna entre la posesión de una creencia tipo (A1) y la posesión de un deseo tipo (A2). Pues en cualquier caso, siempre parece estar abierta la posibilidad de que un sujeto no desee como cree que una versión plenamente racional de sí mismo desearía que lo hiciera.

Para ilustrar esta objeción, nuevamente me remitiré al caso de Miguel planteado en la sección 2 del presente capítulo. En ese caso, tras juzgar moralmente incorrecta su acción de explotar trabajadores, Miguel obtuvo motivación para no consumir productos vendidos por la empresa explotadora. De acuerdo con la teoría racionalista de Smith, la motivación de Miguel se explica en dos pasos. Primero, por el hecho de que al emitir su juicio, expresó su posesión de una creencia según la cual una versión plenamente racional de sí mismo desearía que él, en su versión actual, deseara no comprar productos de la empresa en cuestión. Y segundo, porque dada su creencia acerca de lo que la versión plenamente racional de sí mismo desearía que deseara, Miguel procedió a formar un deseo que se ajusta a esa creencia. El aspecto sobre el cual la objeción de modo de presentación inadecuado llama la atención es

que, en su conjunto, los dos pasos que componen la explicación desde la teoría de Smith no permiten afirmar que Miguel obtuvo motivación por el simple hecho de haber emitido un juicio moral. Más bien, parece que una explicación en semejantes términos sólo podrá dar cuenta de la motivación de Miguel en la medida en que se le atribuya una creencia ulterior con un contenido tal como: "yo quiero desear como desearía mi versión plenamente racional". La razón por la cual la explicación requiere de un estado de este tipo es porque solo así se podría llenar el vacío derivado del uso de un modo de presentación inadecuado en la caracterización del contenido de la creencia que, según la teoría de Smith, fundamenta la motivación de Miguel.

Si esta objeción es correcta, parece que Smith únicamente podrá elegir entre dos alternativas. Una primera alternativa sería optar por preservar los términos de su explicación original. Como he indicado, el costo de elegir esta alternativa sería la renuncia a la posibilidad de afirmar la tesis central del internismo moral, según la cual los juicios morales mantienen una relación interna con la motivación para actuar. Así, la teoría de los juicios morales de Smith únicamente podría funcionar como una solución parcial del problema moral, pues implicaría la negación de una de las proposiciones que lo conforman.

Por otro lado, la segunda alternativa que puede elegir Smith, consiste en remediar la inadecuación de modo de presentación en su explicación introduciendo un recurso teórico ulterior. Una posible estrategia que iría en este sentido consistiría en afirmar, como plantee en lo concerniente al caso de Miguel, que cuando un sujeto obtiene motivación moral, no sólo posee la creencia expresada por su juicio respecto a lo que su versión plenamente racional desea, sino que además posee algún otro estado que permite llenar el vacío que hay entre la posesión de (A1) y la formación de (A2). Por ejemplo, un estado de ese tipo sería la mencionada creencia con un contenido tal como: "yo quiero desear como desearía mi versión plenamente racional". Sin embargo, al reparar en las implicaciones de esta alternativa, podemos percatarnos de que, en la medida en que la explicación pasaría a depender de aquel nuevo recurso teórico introducido, nuevamente la teoría de Smith sería incompatible con el internismo moral. Pues al depender de la mediación de una abstracción como la creencia descrita, contrariamente a lo que distingue al internismo, la teoría estaría afirmando el fetichismo moral.

Al igual que las objeciones anteriores, la objeción de modo de presentación inadecuado ofrece una lección acerca de lo que cabría esperar de una teoría adecuada de los juicios

morales. La lección en cuestión es la siguiente:

> ***Modo de presentación actual:*** una buena teoría de los juicios morales debería de concebir el contenido de esos juicios bajo un modo de presentación que, o bien describa la manera como se concibe actualmente el sujeto que juzga, o bien esté conectado con tal manera de concebirse actualmente, sin introducir subrepticiamente elementos motivadores no explicados.

5. Conclusiones

Este capítulo ha estado dedicado a llevar a cabo dos labores diferentes. Una primera labor ha sido aclaratoria. Específicamente, las secciones 2 y 3 han tenido la función de clarificar tanto qué es lo que se entiende por juicio moral, como cuáles son los elementos propios de la explicación racionalista que Smith ofrece de dicho fenómeno. Mientras que para la primera aclaración procedí a plantear un caso claro de juicio moral, para la segunda, elaboré un esquema en el que se identificaron con precisión los elementos fundamentales de la explicación racionalista. Posteriormente, la segunda labor llevada a cabo durante el capítulo, ha tenido lugar en los tres apartados de la sección 4. Dicha labor ha consistido en desarrollar tres objeciones originales a la teoría racionalista de Smith. La primera objeción se ha centrado en la atribución de un problema de hiper-intelectualización. Por otra parte, la segunda objeción ha mostrado que la teoría de Smith no puede dar cuenta del hecho de que la justificación de la acción que ofrecen los juicios morales es susceptible al cambio. Y finalmente, a través de la tercera objeción he argumentado que, como consecuencia de la introducción de un modo de presentación inadecuado en su explicación de la motivación moral, contrariamente a lo que pretende, la teoría de Smith es incompatible con el internismo moral.

CAPÍTULO III

Una solución sentimentalista al problema moral

1. Introducción

El objetivo de este tercer capítulo es desarrollar una teoría de los juicios morales que pueda funcionar como solución conciliadora al problema moral (PM). En particular, una teoría de los juicios morales que sea mejor solución conciliadora a PM que la solución racionalista ofrecida por Smith. Para desempeñar dicha función, la teoría propuesta tendrá que ser capaz de implicar simultáneamente cada una de las proposiciones que conforman PM, pues es eso lo que permitirá reconocerla como una solución conciliadora a PM. Pero además, para mejorar a la teoría de Smith, la propuesta deberá de poder evitar las tres objeciones planteadas a dicho enfoque racionalista en la sección 4 del capítulo II de este ensayo. La tesis central de la teoría que desarrollaré establece que, si una determinada actitud es un juicio moral, entonces dicha actitud tiene una (o varias) base(s) emocional(es). Dada su afirmación de la idea de que los juicios morales tienen bases emocionales, la teoría se configura como una forma de sentimentalismo moral.[27]

El capítulo se divide en seis secciones. En la sección 2 describiré la concepción general de las emociones a partir de la cual se articula la teoría sentimentalista que propongo. Posteriormente, en la sección 3 introduciré el concepto de emociones morales, el cual me permitirá indicar específicamente qué tipo de emociones son las que considero que ejercen como base de los juicios morales. Tras clarificar qué es lo que entiendo por emoción y más particularmente por emoción moral, en la sección 4 presentaré una formulación final de la teoría que propongo. Más adelante, en la sección 5, mostraré cómo mi teoría, en conjunto con ciertas hipótesis auxiliares, implica cada una de las proposiciones que conforman PM. Y finalmente, en la sección 6, me dedicaré a presentar las razones por las cuales mi teoría sentimentalista no enfrenta las objeciones que afectan a la teoría racionalista de Smith.

[27] Históricamente, el origen de las teorías sentimentalistas sobre los juicios morales suele ubicarse en (Hume, 1978/1739) y (Smith, 2002/1759). En la actualidad, ejemplos de teorías sentimentalistas pueden encontrarse en (Prinz, 2006, 2007) y (Haidt, 2012).

2. Concepción general de las emociones

En filosofía, el debate contemporáneo con respecto a la naturaleza de las emociones gira en torno a la cuestión de cuál es el componente más fundamental de dichos estados mentales.[28] Algunas teorías sostienen que el componente más fundamental de las emociones es su fenomenología, es decir, el tipo de experiencia consciente que involucran.[29] Otras teorías defienden que se trata más bien de su capacidad para desempeñar una función evaluativa de estados de cosas en el mundo.[30] Y de manera análoga, un tercer grupo de teorías afirma que el componente más fundamental de las emociones es su capacidad para motivar la acción.[31] En cualquier caso, partiendo del reconocimiento de las emociones como estados con múltiples componentes, se procede a argumentar a favor del reconocimiento de uno de esos componentes como el más fundamental.

Para el desarrollo de mi propuesta de solución al problema moral (PM) no me comprometeré con ninguna perspectiva particular acerca de cuál es el componente más fundamental de las emociones. Más bien, me limitaré a afirmar una concepción general, compatible con diversas teorías, que les identifica como estados compuestos de cinco componentes diferentes. En concreto, la concepción general que defenderé identifica a las emociones como estados mentales que tienen intencionalidad, que pueden evaluarse según ciertos estándares de corrección, que son diferentes a los juicios, que tienen la capacidad de motivar la acción y que reflejan las preocupaciones de los agentes que los poseen. Semejante concepción general de las emociones permite articular una teoría sentimentalista de los juicios morales capaz de cumplir con los objetivos perseguidos en este capítulo. Es decir, una teoría sentimentalista a partir de la cual se puedan explicar las tres proposiciones que componen PM sin generar los problemas que produce la teoría racionalista planteada por Smith. En los apartados ulteriores de la presente sección me dedicaré a profundizar en la caracterización de cada uno de los cinco componentes de la concepción general de las emociones sobre la cual opera mi teoría.

[28] Para una descripción detallada de este debate, véanse (de Sousa & Scarantino, 2021), (Tappolet, 2023) y (Brady, 2025).
[29] Un ejemplo de las llamadas "teorías del sentimiento de las emociones" es (Goldie, 2000).
[30] Ejemplos de las llamadas "teorías evaluacionistas de las emociones" son (Deonna & Teroni, 2012) y (Tappolet, 2016).
[31] Como ejemplo de las llamadas "teorías motivacionales de las emociones" puede tomarse (Scarantino, 2014).

2.1. Intencionalidad

El primer componente de la concepción general de las emociones subyacente a mi teoría es la intencionalidad. Al igual que otros estados como las creencias y los deseos, las emociones son estados psicológicos intencionales en el sentido de que su contenido es *acerca* de algo. Como cualquier otro estado intencional, aquello acerca de lo cual es el contenido de una emoción es el objeto intencional de dicha emoción. Para constatar el aspecto intencional de las emociones no hace falta más que pensar en diferentes tipos de estados emocionales. Por ejemplo, el miedo ante la presencia de una serpiente constituye un estado que es *acerca* de la presencia de una serpiente, la vergüenza por la calificación obtenida en un examen constituye un estado *acerca* de la calificación obtenida en un examen, la felicidad por un encuentro es un estado *acerca* de un encuentro y la indignación hacia una cierta conducta es un estado *acerca* de una determinada conducta. Independientemente de la posibilidad de que pueda haber algunos estados emocionales que no tengan intencionalidad, dados los objetivos específicos de este trabajo, para el desarrollo de mi concepción general de las emociones bastará con afirmar que en general, las emociones son estados mentales con intencionalidad.

Los objetos intencionales de las emociones pueden ser de diferentes tipos (Tappolet, 2021, pp. 24-26). Sin pretender ofrecer un listado exhaustivo, tan solo retomando los ejemplos previamente descritos, se puede apreciar que una emoción puede tener como objeto intencional estados de cosas como la presencia de una serpiente, objetos abstractos como la calificación de un examen, eventos como el encuentro con una persona o clases de cosas como un cierto tipo de conducta. Además, otro rasgo propio de la intencionalidad de las emociones es el hecho de que aunque no siempre, en muchas ocasiones el objeto intencional de una emoción es, o al menos participa, de la causa de ese misma emoción. Por ejemplo, el miedo ante la presencia de una serpiente tiene su causa en la percepción de una serpiente, la vergüenza por la calificación en un examen tiene su causa en la creencia de que se ha obtenido una determinada calificación en un examen, la alegría por un encuentro tiene su causa en la percepción del encuentro y la indignación ante una acción tiene su causa en la creencia de que la acción relevante ha tenido lugar. En cada caso, la emoción es causada por un estado cognitivo diferente a la emoción pero con el mismo objeto intencional. Dichos estados cognitivos son identificados en la literatura como las bases cognitivas de las emociones (Deonna y Teroni, 2012).

2.2. Condiciones de corrección

El segundo componente de la concepción general de las emociones subyacente a mi teoría de los juicios morales es la posesión de condiciones de corrección. Partiendo de su condición como estados mentales intencionales, las emociones pueden reconocerse también como estados que ofrecen acceso al conocimiento de ciertas propiedades, específicamente, propiedades evaluativas cuya atribución a objetos intencionales explica la activación de una emoción. Por ejemplo, el miedo involucra conocimiento de la propiedad evaluativa de "ser peligroso", la vergüenza supone conocimiento de la propiedad evaluativa de "ser vergonzoso", la alegría implica conocimiento de la propiedad evaluativa de "ser agradable" y la indignación involucra conocimiento de la propiedad evaluativa de "ser indignante". Es el conocimiento de la propiedad evaluativa relevante en cada caso y su atribución a un determinado objeto intencional lo que explica la posesión de emoción de parte de un sujeto. En la literatura, estas propiedades evaluativas que se atribuyen a objetos intencionales son denominadas "objetos formales".[32]

Ahora bien, en la medida en que tener una emoción implica evaluar a un objeto intencional como ejemplificando un objeto formal, una emoción podrá considerarse correcta o incorrecta dependiendo de si dicha evaluación es o no es correcta. Regresando nuevamente a los ejemplos antes descritos, podemos decir que el miedo ante la presencia de una serpiente será correcto si la atribución del objeto formal de "ser peligroso" a la presencia de la serpiente es correcta, la vergüenza por la calificación en un examen será correcta si la atribución del objeto formal de "ser vergonzoso" a la calificación obtenida es correcta, etc. Esta derivación, de condiciones de corrección de las emociones, a partir de su identificación como estados intencionales que involucran acceso epistémico a objetos formales, es descrita con mayor detalle por Echeverri (2019) en su caracterización del enfoque cognitivista acerca de las emociones. De acuerdo con Echeverri (2019, p. 543), partiendo de semejante derivación, la corrección de los estados emocionales en general puede entenderse como una cuestión sujeta a un principio como el siguiente:

[32] La especificación del objeto formal correspondiente a cada emoción es una cuestión abierta. Aquí, me limitaré a decir que el objeto formal de una emoción es aquello a lo cual la emoción en cuestión es sensible.

Corrección de emociones (CE): "una emoción *E* es correcta si y sólo si su objeto intencional, *O*, ejemplifica el objeto formal que *E* (re)presenta a *O* como teniendo. Es incorrecta en cualquier otro caso. (Echeverri, 2019, p. 543)"[33]

2.3. Diferencia respecto de los juicios

El tercer componente de la concepción general de las emociones subyacente a mi teoría de los juicios morales es el contraste con respecto a los juicios. Considerar a las emociones como estados mentales intencionales y con condiciones de corrección no implica su identificación con juicios.[34] Una razón para distinguir entre emociones y juicios es la existencia de casos de emociones recalcitrantes. De acuerdo con el uso del término en la literatura, una emoción recalcitrante es una emoción tal que persiste en la psicología de un sujeto aún cuando ese mismo sujeto sostiene un juicio incompatible con la corrección de dicha emoción. Es decir, en un caso de emoción recalcitrante, un sujeto juzga que un objeto intencional *O* no ejemplifica una determinada propiedad evaluativa, y sin embargo, posee una emoción que atribuye a *O* esa misma propiedad evaluativa.

Una emoción que suele manifestar recalcitrancia es el miedo. Pensemos en un sujeto *S* que tiene miedo de viajar en avión, una emoción que para ser correcta requiere de que el objeto intencional de viajar en avión ejemplifique el objeto formal de "ser peligroso". La idea es que el miedo de *S* será recalcitrante si persiste aún cuando, contrariamente a lo que tendría que ser el caso para que su miedo sea correcto, *S* juzga que viajar en avión en realidad no es peligroso. Por ejemplo, si persiste aún cuando *S* juzga que viajar en avión no es peligroso porque las probabilidades de que un avión se accidente son de aproximadamente 1 entre 11 millones.

Las emociones recalcitrantes muestran que las emociones no son juicios por el simple hecho de que dan pie a casos en los que las emociones no se comportan como juicios. En concreto, las emociones recalcitrantes dan pie a casos en los que una emoción no manifiesta rasgos propios de los juicios tales como la sensibilidad a la evidencia o la integración

[33] El uso de "(re)presentación" en *CE* busca llamar la atención sobre el hecho de que la manera en que las emociones atribuyen objetos formales a objetos intencionales puede entenderse tanto en términos representacionales como no-representacionales. En términos representacionales, una emoción atribuye un objeto formal a un objeto intencional a través de la representación del objeto intencional ejemplificando al objeto formal. En cambio, en términos no-representacionales, una emoción atribuye un objeto formal a un objeto intencional a través de la activación de disposiciones para actuar asociadas con el objeto formal relevante.

[34] Algunas teorías cognitivistas como la de Solomon (1980) y la de Nussbaum (2001) defienden semejante identificación entre emociones y juicios.

inferencial. Pues en casos como el previamente descrito de un sujeto con miedo a volar, con lo que nos encontramos es con una emoción que prevalece aún cuando hay razones disponibles que indican su incorrección, y que además, no se integra inferencialmente en el sentido de que no le afecta estar en contradicción con otros estados.

2.4. Motivación de la acción

El cuarto componente de la concepción general de las emociones subyacente a mi teoría de los juicios morales es la capacidad de motivar la acción. Al evaluar objetos intencionales a través de objetos formales, las emociones generan un impacto sobre la motivación que se refleja en disposiciones a actuar en concordancia con la evaluación relevante. Por ejemplo, al evaluar un objeto intencional como peligroso, el miedo genera un impacto sobre la motivación que se refleja en la disposición a evitar dicho objeto intencional. Y de manera equivalente, al evaluar un objeto intencional como culposo, la culpa genera un impacto sobre la motivación que se refleja en la disposición a remediar el daño producido por el objeto intencional en cuestión. A manera de síntesis, el componente motivacional de las emociones puede plasmarse en el siguiente principio:

> ***Motivación emocional (ME):*** si *E* es una emoción, entonces *E* conlleva una motivación para actuar conforme a su objeto formal *f*.

La idea de fondo en la configuración de este componente es, pues, que en la medida en que una emoción involucra la evaluación de un objeto intencional, dicha emoción genera un impacto sobre la motivación que se refleja en la activación de disposiciones a actuar de formas específicas. Si bien algunas emociones como ciertos tipos de tristeza podrían parecer excepciones en el sentido de que no se asocian con acciones concretas, su supuesta excepcionalidad puede cuestionarse a través de la idea de que en realidad constituyen casos de emociones que motivan la inacción.[35]

En la literatura, un argumento especialmente influyente para respaldar el componente motivacional de las emociones es desarrollado tanto por Frijda (1986) como por Deonna y

[35] Mientras que autores como (Prinz, 2004) y (Tappolet, 2010; 2016) argumentan que casos como la tristeza por duelo sí constituyen contraejemplos a la tesis de que las emociones en general motivan la acción, autores como (Scarantino, 2014) responden a dichos argumentos defendiendo la idea de que la particularidad de semejantes estados es que motivan la inacción.

Teroni (2012). Dicho argumento se centra en la fenomenología de los estados emocionales, es decir, en el tipo de experiencia que atravesamos al tener una emoción. La idea es que cuando experimentamos una emoción, la posibilidad de reconocer esa experiencia como la de una emoción en particular depende de manera fundamental de la manifestación, como parte de la experiencia, del tipo de motivación que asociamos con la emoción relevante. Por ejemplo, cuando experimento ira frente a la acción de mi vecino de subir el volumen de la música a mitad de la noche, reconozco que mi experiencia es de ira porque, como parte de la experiencia que tengo, encuentro motivación para ir a enfrentar a mi vecino y obligarlo a bajar el volumen de la música. Tanto la ausencia de motivación para llevar a cabo una acción semejante, como el reconocimiento de una disposición a actuar de una manera completamente diferente, me darían razón para pensar que la experiencia que tengo en realidad no es una experiencia de ira. De hecho, sin una disposición semejante, probablemente ni siquiera me plantearía la posibilidad de que estoy experimentando ira. El mismo argumento puede formularse apelando a cualquier otra emoción. Si por ejemplo, me siento culpable por haber robado un par de productos del supermercado, reconozco mi experiencia como una de culpa porque identifico una disposición para ir a regresar los productos robados. Sin una disposición semejante, tendría razón para pensar que en realidad no siento culpa o ni siquiera me plantearía la posibilidad de que esté experimentando culpa. En resumen, el argumento respalda la afirmación del componente motivacional de las emociones en la medida en que deja ver que la motivación se articula como un elemento intrínseco a la experiencia de una emoción.

2.5. Reflejos de preocupaciones

Finalmente, el quinto y último componente de la concepción general de las emociones subyacente a mi teoría de los juicios morales es la capacidad de reflejar las preocupaciones de quien las posee. Para especificar qué es lo que supone este componente, resulta imprescindible definir el concepto de preocupación. Una definición influyente en la literatura sobre emociones es elaborada por Roberts (2003, pp. 141-151). De acuerdo con Roberts, las preocupaciones son estados mentales en la psicología de un sujeto que permiten explicar las respuestas emotivas de ese mismo sujeto. Conocer que un determinado sujeto *S* posee una preocupación en particular, permite inferir un conjunto de posibles respuestas emotivas que

cabría esperar de *S*. La preocupación se articula, pues, como razón para que *S* adopte ciertas respuestas emotivas y no otras.

Desde la perspectiva de Roberts, las preocupaciones no son estados mentales de un cierto tipo, sino que diferentes tipos de estados mentales pueden erigirse como una preocupación. Por ejemplo, los deseos son estados que con regularidad se configuran como preocupaciones. Si *S* desea comer, entonces cabrá esperar que *S* tenga emociones como enojo o frustración mientras su deseo de comer no se satisfaga. O si *S* desea que se le reconozca como un buen filósofo, entonces cabrá esperar que *S* tenga emociones como orgullo o felicidad si uno de sus artículos es aceptado para publicación en una revista de renombre. Sin embargo, en la medida en que no todas las preocupaciones involucran compromisos con objetivos, no todas las preocupaciones constituyen deseos. Estados como los apegos y las aversiones funcionan también como preocupaciones. Por ejemplo, el apego de un padre hacia su hija se articula como una de sus preocupaciones más fundamentales, en el sentido de que no solo fundamenta respuestas emocionales por sí mismo, sino que también deriva preocupaciones ulteriores como el deseo de que su hija tenga salud. Asimismo, la aversión de una persona hacia los insectos no solo fundamenta respuestas emocionales como el miedo frente a la presencia de una araña, sino que también fundamenta preocupaciones derivadas como la aversión ulterior a hacer caminatas en el campo.

Una preocupación puede ser ocurrente o meramente disposicional. Al ser del primer tipo, la preocupación es tal que el sujeto que la posee atiende conscientemente a ella en la determinación de qué tipo de reacciones emocionales serían apropiadas dadas ciertas circunstancias. En cambio, al ser disposicional, la preocupación ejerce su influencia sobre las respuestas emocionales del sujeto sin que este se percate directamente de ello. La misma preocupación puede ser ocurrente en un tiempo *T1* y disposicional en un tiempo diferente *T2*. Por otra parte, Roberts (2003, p.142) advierte que las preocupaciones pueden distinguirse también según su origen. En concreto, distingue entre preocupaciones biológicas o instintivas y preocupaciones aprendidas. Como ejemplo de las primeras, cita a la preocupación por evitar el daño físico y a la preocupación por evitar la muerte. Y para ilustrar a las segundas, se remite a la preocupación por evitar el cáncer y a la preocupación por que la calidad del trabajo que uno realiza sea reconocida.

Independientemente del tipo de estado mental que sean, de su grado de fundamentalidad, de si hay o no consciencia respecto a su influencia y de si son instintivas o aprendidas, las preocupaciones en general constituyen disposiciones a responder emocionalmente de ciertas maneras. Y puesto que las emociones son estados con un

componente motivacional, las preocupaciones pueden entenderse también como disposiciones a encontrar motivación para actuar de ciertas maneras. En su conjunto, las disposiciones de un sujeto, derivadas de sus preocupaciones específicas, a responder emocional y conductualmente de ciertas maneras, constituyen el carácter o personalidad de ese sujeto en particular. Pues conforme dichas disposiciones marcan patrones, delimitan una forma de ser específica que permite caracterizar al sujeto que las posee.

3. Emociones morales

Tras haber presentado una concepción general de las emociones, procederé a introducir la categoría más específica de las emociones morales. La importancia de introducir esta categoría es consecuencia de que son las emociones morales en particular las que desempeñan una función dentro de mi teoría sentimentalista de los juicios morales. En concreto, son las emociones morales, en tanto estados con los cinco componentes atribuidos a las emociones en general, pero que además manifiestan otros dos rasgos característicos, el tipo de estados mentales a los que identifico como base de los juicios morales.

En la literatura en psicología moral y metaética, una definición de las emociones morales que ha sido especialmente influyente es elaborada por Jonathan Haidt (2003). De acuerdo con dicha definición, las emociones morales son "aquellas emociones que están vinculadas a los intereses y bienestar de la sociedad como un todo o al menos de otras personas diferentes al agente que las posee" (Haidt, 2003, p.853). Partiendo de esta caracterización general, Haidt procede a identificar dos rasgos prototípicos de las emociones morales. El primero de estos rasgos es la posibilidad de activación a partir de elicitadores desinteresados. Las emociones morales pueden activarse a partir de la interacción con objetos intencionales que no son de relevancia directa para el agente que las posee. Por ejemplo, enterarme de que el presidente de un país lejano al mío es corrupto o ver una fotografía antigua de un niño hambriento son objetos intencionales que pueden activar en mí una emoción moral. Y el segundo rasgo prototípico de las emociones morales es su manifestación de un componente motivacional con una tendencia prosocial. Como cualquier otra emoción, las emociones morales se relacionan necesariamente con la motivación para la acción. Sin embargo, en contraste con la motivación de muchas emociones no-morales, la motivación de emociones morales promueve la satisfacción de intereses sociales y no individuales. Por ejemplo, una emoción moral puede llevarme a tener motivación para actuar de una manera

que sea perjudicial para mis intereses individuales, en el sentido de que me implique un castigo, pero benéfica para los intereses colectivos, en el sentido de que permita remediar un daño.

Partiendo de la identificación de las emociones morales en general como aquellas que presentan los dos rasgos prototípicos descritos en el párrafo anterior, Haidt pasa a distinguir entre cuatro familias de emociones morales. La primera familia de emociones morales es la de las *emociones condenatorias del otro*, las cuales se articulan como respuestas a la identificación de violaciones a normas sociales de parte de otros sujetos diferentes al que posee la emoción. En esta primera familia, Haidt ubica a la ira, a la indignación y al desprecio. Por otra parte, la segunda familia de emociones morales según la taxonomía de Haidt es la de las *emociones de autoconciencia negativa*, las cuales se distinguen por presentarse en un sujeto cuando éste evalúa su propia conducta como deficiente frente a un determinado estándar normativo. Aquí, Haidt ubica a la culpa y a la vergüenza. La tercera familia de emociones morales son las *emociones de reconocimiento del sufrimiento de otro*, las cuales se caracterizan por activarse ante la percepción de dolor o vulnerabilidad ajena. En esta familia, se encuentran como integrantes la compasión y la simpatía. Y finalmente, la cuarta familia de emociones morales de acuerdo con Haidt corresponde a la de las *emociones de elogio del otro*, las cuales se dan como respuesta a acciones o rasgos reconocidos como correctos o valiosos según algún estándar normativo socialmente relevante. Las emociones que participan de esta cuarta familia son la gratitud y la admiración.

4. Las emociones morales como bases de los juicios morales

Recordando lo dicho en la introducción al presente capítulo, la tesis central de la teoría sentimentalista de los juicios morales que propongo como solución a PM es que si una determinada actitud es un juicio moral, entonces dicha actitud tiene una (o varias) base(s) emocional(es). Como indiqué al inicio de la sección 3, la formulación de esta tesis puede precisarse mediante la especificación del hecho de que el tipo de emociones que actúan como bases de los juicios morales son las denominadas "emociones morales". Así, tomando en cuenta este hecho, lo que la tesis contempla es que si una determinada actitud es un juicio moral, entonces dicha actitud tendrá bases en emociones morales. Esta reformulación de la tesis resulta importante en dos sentidos diferentes. Primero, porque permite esclarecer qué tipo de fundamento normativo es el que la teoría atribuye a los juicios morales. En concreto,

se trata de un fundamento cimentado en la sensibilidad de las emociones morales hacia los intereses de la sociedad en su conjunto. Y en segundo lugar, la reformulación es importante porque permite restringir el alcance de la concepción general de las emociones descrita a lo largo de la sección 2 de este capítulo. Específicamente, la reformulación me da oportunidad de aclarar que para que mi teoría sea verdadera, lo único que tiene que ser el caso es que, al menos en lo particular de las emociones morales, la atribución de los cinco componentes antes mencionados sea correcta. Es decir, basta con que las emociones morales sean estados con intencionalidad, que poseen condiciones de corrección, que son diferentes a los juicios, que motivan la acción y que reflejan las preocupaciones de quien las posee.

Pero la apelación a las emociones morales no es la única forma en la que la tesis de la teoría sentimentalista que defiendo puede precisarse. Una cuestión ulterior que también permite precisar es la concerniente al significado del concepto de "base". En este respecto, considero pertinente señalar que el concepto de base que utilizo en la formulación de la tesis de mi teoría es entendido en apego al siguiente principio:

> ***Base:*** un estado *E* es base de una actitud *A*, si y sólo si, *E* es aquello que explica porqué un sujeto adopta la actitud *A*, si acaso la adopta.

Tomar en consideración *Base* habilita una nueva reformulación de la tesis de mi teoría sentimentalista. De acuerdo con esta nueva reformulación, lo que la tesis contempla es que si una actitud es un juicio moral, entonces, la adopción de dicha actitud de parte de un sujeto *S* se explica por la posesión, de parte de *S*, de emociones morales. Semejante reformulación no solo termina de clarificar qué es exactamente aquello con lo que se compromete mi teoría sentimentalista, sino que además permite llamar la atención sobre un aspecto que distingue a esta teoría de la teoría racionalista planteada por Smith. Se trata del hecho de que mientras la teoría de Smith realiza una afirmación con respecto a qué es lo que son los juicios morales, mi teoría apunta más bien a especificar qué es lo que explica la realización de juicios morales. Así, mi teoría no se compromete con una afirmación tipo "los juicios morales son *X*", sino con una afirmación tipo "los juicios morales se explican por *X*".[36]

[36] Desde una perspectiva funcionalista, la pregunta del primer tipo sólo podría responderse a través de una respuesta a la pregunta del segundo tipo. Esto se debe a que desde el funcionalismo, la naturaleza de una actitud se determina por las relaciones causales que sostiene con otras entidades.

5. El sentimentalismo moral como solución al problema moral

Durante las siguientes tres secciones, me dedicaré a mostrar cómo es que mi teoría sentimentalista de los juicios morales permite explicar cada una de las proposiciones que, según Smith (1994), conforman al problema moral (PM). En última instancia, si la teoría efectivamente es capaz de dar cuenta de las tres proposiciones en cuestión, entonces se le podrá reconocer como una solución conciliadora a PM. Es decir, como una solución capaz de resolver el problema sin necesidad de negar la verdad de alguna de las proposiciones involucradas en su formulación.

5.1. Sentimentalismo y cognitivismo moral

La primera proposición que tendría que implicarse a partir de una teoría de los juicios morales capaz de funcionar como una solución conciliadora a PM establece que: "los juicios morales de la forma "es correcto que yo haga p", expresan las creencias de una persona sobre una cuestión objetiva de hecho, un hecho acerca de lo que es correcto hacer." (Smith, 1994, p. 12). Esta proposición captura la tesis central del cognitivismo moral, según la cual los juicios morales son estados mentales cognitivos, en el sentido de que poseen condiciones de corrección que dependen de su capacidad para referir a hechos objetivos.

La teoría sentimentalista de los juicios morales que propongo como solución a PM es capaz de implicar la verdad de esta primera proposición una vez que se le considera en conjunción con una determinada hipótesis auxiliar. De acuerdo con dicha hipótesis, si un estado antecedente hace inteligible un estado ulterior, será razonable esperar que las condiciones de corrección del segundo dependan de las condiciones de corrección del primero. Para mostrar cómo es que la teoría, en conjunto con esta hipótesis auxiliar, implica al cognitivismo moral, es necesario aclarar qué es lo que contempla la hipótesis y porqué resulta plausible.

La idea de un estado mental antecedente que hace inteligible a un estado mental ulterior puede ilustrarse apelando al tipo de relación que muchas veces se da entre percepciones y creencias. Por ejemplo, si cuando veo una manzana frente a mí procedo a formar la creencia de que hay una manzana enfrente de mí, el tipo de relación que hay entre mi percepción y mi creencia es una en la que la primera hace inteligible a la segunda. La mencionada inteligibilidad simplemente refiere al hecho de que es la percepción lo que

explica por qué he formado la creencia. Si procediera a formar la creencia sin una percepción correspondiente, difícilmente podría entender por qué la he formado. De hecho, mantener tal creencia no parecería razonable. Otro caso común de relación de inteligibilidad entre estados mentales se da entre diferentes creencias dentro de contextos inferenciales. Por ejemplo, cuando dos creencias en condición de premisas dan lugar, a manera de conclusión, a una creencia ulterior. Si creo que hoy va a llover y creo, además, que si llueve cancelarán el partido de fútbol, entonces procedo a creer que el partido de fútbol será cancelado. Nuevamente, las dos creencias iniciales hacen inteligible a la tercera en el sentido de que permiten explicar por qué la he formado. Predicada del vínculo entre emociones morales y juicios morales, la relación de inteligibilidad supone exactamente lo mismo. Por ejemplo, si cuando siento indignación frente a la acción de mi primo de patear a su perro procedo a juzgar la acción de mi primo como moralmente incorrecta, el tipo de relación que hay entre mi emoción y mi juicio moral es una en la que la primera hace inteligible al segundo. La emoción hace inteligible al juicio en el sentido de que es ella lo que me permite explicar por qué he emitido el juicio. Sin la emoción, difícilmente podría explicar por qué he emitido un juicio semejante.

Ahora bien, la importancia de la relación de inteligibilidad para comprender cómo es que mi teoría sentimentalista implica al cognitivismo moral se debe a la capacidad de esta relación para indicar cómo es que las condiciones de corrección de un estado pueden depender de las condiciones de corrección de otro estado. En concreto, lo que esta relación permite mostrar es que cuando un estado antecedente hace inteligible a otro estado ulterior, resulta razonable esperar que las condiciones de corrección del segundo dependan de las condiciones de corrección del primero. Para ilustrar la cuestión podemos remitirnos nuevamente al vínculo entre una percepción y una creencia. Por ejemplo, si mi percepción de una manzana frente a mí es verdadera y hace inteligible a mi creencia ulterior de que hay una manzana enfrente de mí, resulta razonable suponer que mi creencia también será verdadera. Por el contrario, si la percepción fuera falsa, en la medida en que hace inteligible a la creencia, lo que cabrá esperar es que la creencia también sea falsa. Lo mismo sucede en el caso del par de creencias que, en tanto premisas, permiten inferir y en consecuencia hacer inteligible a una creencia ulterior. Si las premisas son verdaderas, entonces, lo que cabrá esperar es que la conclusión también lo sea. Y si en cambio, las premisas son falsas, entonces no habrá razón para esperar que la conclusión sea verdadera. En lo concerniente al vínculo entre una emoción moral y un juicio moral la cuestión se articula en los mismos términos. Si mi indignación frente a la acción de mi primo es correcta y hace inteligible a mi juzgar

moralmente incorrecta a esa misma acción, entonces, resulta razonable suponer que mi juicio moral también será correcto. Y por el contrario, si mi emoción es incorrecta, entonces, lo que cabrá esperar es que mi juicio moral también sea incorrecto.

A partir de la clarificación anterior de la hipótesis auxiliar de inteligibilidad, puede apreciarse cómo es que dicha hipótesis, en conjunción con mi teoría sentimentalista, puede implicar el cognitivismo moral. La teoría contribuye con dos afirmaciones diferentes. En primer lugar, sostiene que las emociones morales, en tanto base, explican y en consecuencia hacen inteligibles a los juicios morales. Y en segundo lugar, afirma (como se indicó en las secciones 2.1 y 2.2 del presente capítulo) que las emociones morales son estados cognitivos, que poseen intencionalidad y tienen condiciones de corrección que dependen de su capacidad para referir a hechos objetivos. En concreto, recordando la teoría de las emociones morales de Haidt (2003), hechos objetivos con respecto a los intereses y el bienestar de la sociedad como un todo o de otras personas diferentes a la que posee la emoción relevante.[37] Por su parte, la hipótesis auxiliar muestra que si un estado antecedente hace inteligible a otro, entonces, cabe esperar que las condiciones de corrección del segundo dependan de las del primero. Así, queda claro cómo es que si tanto la teoría como la hipótesis están en lo cierto, entonces puede concluirse que los juicios morales poseen condiciones de corrección que dependen de la referencia a hechos objetivos.

5.2. Sentimentalismo e internismo moral

La proposición (ii) de PM señala que: "si alguien juzga que es correcto hacer *p*, entonces, *ceteris paribus*, está motivado a hacer *p*" (Smith, 1994, p. 12). De acuerdo con Smith, esta proposición corresponde con una cierta forma de internismo moral. En general, el internismo constituye un enfoque comprometido con la afirmación de una conexión entre emitir un juicio moral y obtener motivación para actuar que tiene un fundamento interno a los juicios morales. Pero recordando lo dicho en el capítulo I de esta tesis, el internismo en (ii) se caracteriza por dos aspectos. En primer lugar, es un internismo débil, en el sentido que reconoce y acomoda la existencia de casos en los que, pese a juzgar moralmente, un sujeto no obtiene motivación para actuar. Y en segundo lugar, es un internismo capaz de explicar la

[37] Más allá de la caracterización específica que ofrece la teoría de Haidt, el aspecto relevante es que mi teoría sentimentalista es compatible con concepciones objetivistas de los hechos morales.

motivación moral sin afirmar el fetichismo moral. Si mi teoría sentimentalista puede explicar un internismo moral semejante, entonces podrá explicar (ii).

La posibilidad de explicar esta segunda proposición de PM desde mi teoría sentimentalista depende de la posibilidad de mostrar que las emociones morales, en su condición de bases de los juicios morales, son estados mentales que sostienen una conexión interna con la motivación para actuar. Pues si lo que explica, en tanto base, un juicio moral es la posesión de una emoción moral, entonces, si un sujeto juzga moralmente posee una emoción moral. Y si las emociones morales son estados que mantienen una relación interna con la motivación para actuar, entonces, se puede concluir que los juicios morales mantienen una relación interna con la motivación para actuar.

Ahora bien, para mostrar que las emociones morales son estados mentales que necesariamente implican motivación, nuevamente recurriré a la postulación de una hipótesis auxiliar. En este caso, la hipótesis relevante contempla que: si una entidad *E1* mantiene una relación interna con una entidad *E2*, entonces, la ausencia de *E2* da razón para pensar en la ausencia de *E1*. En última instancia, a través de esta hipótesis argumentaré que, en tanto la ausencia de motivación para actuar de razón para pensar en la ausencia de una emoción moral, puede afirmarse que las emociones morales mantienen una relación interna con la motivación. Sin embargo, antes de semejante aplicación de la hipótesis relevante es necesario mostrar su plausibilidad. Para ello, basta con pensar en casos de entidades que sostengan una relación interna con alguna otra entidad. Por ejemplo, pensemos en la relación entre ser un triángulo y ser una figura con tres ángulos. Si en el contexto de una clase de matemáticas el profesor me pregunta si hay un triángulo en el pizarrón y tras ver el pizarrón me percato de la ausencia de una figura con tres ángulos, evidentemente, esa información me dará razón para pensar que no hay un triángulo en el pizarrón. Como indica la hipótesis, la ausencia en el pizarrón de una figura con tres ángulos indica que no hay un triángulo en el pizarrón y eso, confirma, que los triángulos son entidades que mantienen una relación interna con la propiedad de ser una figura con tres ángulos. Otro ejemplo que respalda la plausibilidad de la hipótesis auxiliar relevante concierne a la relación entre ser padre y tener hijos. Si al conocer a una persona esta me comenta que no tiene hijos, dicha información me dará una razón bastante sólida para pensar que la persona que acabo de conocer no es padre. Pero nuevamente, más allá de ese dato acerca de mi interlocutor, lo que la experiencia muestra es que hay una relación interna entre ser padre y tener hijos.

En lo particular de la relación entre emociones y motivación para actuar, la hipótesis auxiliar relevante es ejemplificada en el contexto de un argumento que ya he presentado. Me refiero al argumento fenomenológico desplegado tanto por Frijda (1986) como por Deonna y Teroni (2012), precisamente para defender la identificación de la motivación como un componente intrínseco a las emociones en general. Como señalé en la sección 2.4, el punto de partida del argumento se encuentra en el hecho de que cuando experimentamos una emoción, la posibilidad de reconocer la experiencia que tenemos como la de una emoción en particular, depende de la manifestación en dicha experiencia de motivación para actuar de formas asociadas con la emoción relevante. Partiendo de ese hecho, el argumento pasa a afirmar que dado el rol imprescindible de la motivación para determinar si tenemos o no una emoción, puede concluirse que las emociones mantienen una relación interna con la motivación para actuar. El argumento puede ejemplificarse a través de un caso como el siguiente. Imaginemos un sujeto *S1* que al salir de un restaurante de una cena con su pareja encuentra un perro callejero evidentemente hambriento. Tras percatarse de la presencia del animal, *S1* le comenta a su pareja que siente compasión por él. Sin embargo, pese a tener en su mano una bolsa llena de restos que no se terminaron durante la cena, *S1* no encuentra ninguna motivación para al menos regalar un poco de comida al perro. Partiendo de la ausencia de motivación en su experiencia, *S1* reconoce que en realidad no ha sentido compasión, sino que más bien, al realizar el comentario, se estaba intentando mostrar como una persona sensible a los ojos de su pareja. De manera similar, un segundo ejemplo que también permite apreciar cómo opera el argumento fenomenológico involucra a la emoción de culpa. Imaginemos un sujeto *S2* que realiza comentarios racistas con respecto a uno de sus compañeros de trabajo *S3*. Pensando que nadie en la oficina escuchaba, durante una llamada telefónica *S2* utiliza calificativos como "simio" y "esclavo" para referirse a *S3*. Sin embargo, contrariamente a lo que *S2* suponía, *S3* aún se encontraba en el lugar y escuchó los comentarios. Pese a percatarse de la presencia de *S3*, *S2* simplemente pasa de largo y sale de la oficina. Al día siguiente, después de que *S3* le describiera lo sucedido al jefe de la oficina, *S2* es llamado para dar explicaciones. En ese contexto, *S2* manifiesta sentirse culpable y pide una disculpa a *S3*. Sin embargo, pese a su manifestación de culpa, la ausencia de interés por disculparse en el momento en el que fue descubierto por *S3* y en general de llevar a cabo alguna acción que remediara el daño, permite que *S2* se percate de que en realidad no siente culpa.

Más allá de su confirmación a través de casos de auto-atribución en primera persona, la plausibilidad de la hipótesis auxiliar relevante en lo concerniente a la relación entre emociones morales y motivación puede confirmarse también desde una perspectiva externa.

Para ver cómo esto es así, podemos reformular los dos ejemplos previamente descritos. Primero, pensemos en el caso que refiere a la emoción de compasión. En ese caso, la hipótesis auxiliar puede confirmarse a través del hecho de que, la ausencia de motivación de parte de *S1* para dar comida al perro, no solo ha llevado a que al mismo *S1* piense que en realidad no siente compasión, sino que también ha derivado en que la pareja de *S1* piense que la declaración de compasión no ha sido genuina. Algo muy similar sucede con el caso protagonizado por *S2*. Ahí, el hecho de que al encontrarse descubierto por *S3*, *S2* no se disculpara ni manifestara ningún tipo de motivación para hacer algo que remediara el daño generado por sus comentarios racistas, no solo permite el mismo *S2* sea consciente de que en realidad no siente culpa, sino que también da razón a *S3* para suponer que en realidad *S2* no se siente culpable.

Partiendo del conjunto de ejemplos previamente descritos, puede apreciarse cómo es que mi teoría sentimentalista de los juicios morales, en conjunto con la hipótesis auxiliar relevante, puede explicar la tesis internista según la cual los juicios morales sostienen una conexión interna con la motivación para actuar. Al afirmar que las emociones morales son bases de los juicios morales, la teoría muestra que realizar un juicio moral implica poseer una emoción moral. Y por otra parte, al mostrar que las emociones morales son estados que mantienen una relación interna con la motivación para actuar, la hipótesis auxiliar respalda la afirmación de que la emisión de juicios morales sostiene una relación interna con la motivación. Sin embargo, para mostrar que mi teoría efectivamente puede explicar la proposición (ii) de PM en plenitud, es necesario señalar en qué sentido el internismo que la teoría explica es débil y, por otro lado, la razón por la cual la explicación de la motivación moral resultante no implica afirmar el fetichismo moral.

Primero, el tipo de internismo propuesto es un internismo débil en el sentido de que la relación interna afirmada entre juicio moral y motivación puede tener excepciones. Específicamente, el alcance de dicha relación se restringe a lo particular de sujetos sensibles a emociones morales. La cuestión puede ilustrarse haciendo referencia a casos específicos como los de psicopatía. Mientras que desde la teoría de Smith un psicópata es un sujeto irracional, desde mi teoría, un psicópata es un sujeto que carece de un cierto tipo de sensibilidad emocional. En concreto, de acuerdo con mi teoría, un psicópata es un sujeto que no tiene sensibilidad a las emociones morales en general.

Por otro lado, con respecto a la capacidad de la teoría para no afirmar el fetichismo moral, la cuestión se resuelve en el hecho de que el tipo de motivación que genera una

emoción moral no es mediada por ningún tipo de abstracción. Esta afirmación encuentra respaldo en el hecho de que, como indica Scarantino (2014, p.157), en general, "las acciones emocionales más prototípicas son las acciones intencionales realizadas *al calor del momento*". Pero además, se trata de algo que podemos constatar en la existencia de casos de sujetos que si bien poseen una emoción moral, carecen de los conceptos requeridos para especificar esa emoción moral o atribuirla lingüísticamente. Esto es relevante pues permite mostrar que la motivación que genera una emoción moral ni siquiera requiere de que el sujeto que la experimenta sepa que dicha motivación es reflejo de su posesión de un estado en particular.

5.3. Sentimentalismo y la teoría humeana de la motivación

La proposición (iii) de PM establece que: "un agente está motivado a actuar de cierto modo si tiene un deseo apropiado y una creencia medio-fin, donde creencias y deseos son, en la terminología de Hume, existencias distintas." (Smith, 1994, p.12). Esta proposición captura lo que en la literatura se conoce como la teoría humeana de la motivación (THM).

Un primer compromiso inherente a cualquier versión de THM es la afirmación según la cual creencias y deseos son estados necesariamente diferentes. Tanto la teoría racionalista de Smith como mi teoría sentimentalista son enfoques compatibles con dicha afirmación. En lo particular de mi teoría, esta compatibilidad se hace patente una vez que se aprecia el hecho de que, aunque la teoría afirma que tanto creencias como deseos pueden tener su origen en una emoción, de ello no se sigue que estos estados no sean diferentes. Para confirmar lo anterior no hace falta más que percatarse de que en general, podemos distinguir entre entidades que comparten una causa común. Pensemos en la manera en que, por ejemplo, los distintos hijos de una misma pareja no dejan de ser diferentes por tener una causa compartida, o en la forma en que una infinidad de gotas de agua son causadas por una misma nube sin que ello implique que no son diferentes unas de las otras. Aunque triviales, estos ejemplos permiten advertir que sería equivocado suponer que de la afirmación de que tanto creencias como deseos pueden tener su causa en emociones, se sigue que estos dos tipos de estados no son diferentes.

Ahora bien, más allá de la afirmación de que creencias y deseos son necesariamente diferentes, THM se articula como una teoría comprometida con otras dos tesis, una más débil

que la otra. La tesis débil señala que la posesión tanto de una creencia como de un deseo es una condición *necesaria* para tener motivación. Y la tesis fuerte contempla que la posesión tanto de una creencia como de un deseo es una condición *suficiente* para tener motivación. La teoría racionalista de Smith es compatible con la verdad de estas dos tesis. En cambio, puesto que para dar cuenta de al menos algunos casos de motivación, sostiene que la posesión de una creencia y un deseo no es suficiente sino que también se requiere de una emoción, mi teoría sentimentalista únicamente es compatible con la tesis débil. La compatibilidad de mi teoría con la tesis débil de THM se fundamenta en el hecho de que nada en ella impide afirmar que la explicación de la motivación requiere de la referencia a creencias y deseos. Es decir, nada en mi teoría sentimentalista contradice la afirmación de que creencias y deseos son necesarios para explicar la motivación.

Asumiendo, con base en lo dicho hasta ahora, la compatibilidad de mi teoría sentimentalista con la tesis débil de THM, podríamos proceder a plantear la pregunta de si además, mi teoría podría implicar una tesis semejante. La respuesta a esta pregunta dependerá de los términos en los que se entienda la relación entre las emociones y la acción, algo que exige ir más allá de la concepción general de las emociones que he formulado para plantear mi teoría. Sin embargo, al menos en primera instancia, parece que la mencionada relación de implicación resulta plausible. La razón por la cual esto es así es porque podemos pensar en casos en los que la explicación de la motivación a través de emociones parece requerir de la referencia a creencias y deseos. A continuación, procederé a describir casos de ese tipo.

Para ilustrar el sentido en el cual la referencia a una creencia es requerida en el contexto de una explicación emocional de la motivación, pensemos en el caso de un sujeto *S* que siente una gran indignación hacia un líder político. La indignación de *S* responde al hecho de que el político en cuestión es conocido por ser sumamente corrupto y lograr mantenerse impune. Dada su indignación, *S* siempre ha deseado tener la oportunidad de encarar al político para reclamarle por sus acciones. Sin embargo, aunque en el presente *S* se encuentra comiendo en el mismo restaurante que el político, *S* no tiene motivación para proceder a encararlo. La razón por la cual *S* no tiene motivación es simplemente porque no posee la creencia de que el político se encuentra a tan solo unos metros de distancia. Es decir, aunque posee una emoción y un deseo, *S* no posee la creencia adecuada para tener motivación.

Por otro lado, para ilustrar el sentido en el cual la referencia a un deseo es requerida en el contexto de una explicación emocional de la motivación podemos pensar en el caso de

un sujeto *S1* que siente gratitud. Tras sufrir un accidente de coche, *S1* perdió la consciencia y su auto comenzó a incendiarse. Al percatarse de la situación, una persona *S2* que pasaba por el lugar decidió arriesgar su vida y acercarse al auto en llamas para sacar a *S1*. Después de recuperar la consciencia y enterarse de lo sucedido a través del testimonio de diversos testigos, *S1* se sintió profundamente agradecido con *S2*. En consecuencia, *S1* decidió compensar a *S2* regalándole la totalidad de sus ahorros. Si bien en un caso semejante la creencia de que *S2* lo ha salvado y la gratitud que siente permiten explicar la motivación de S1 para compensar a *S2*, es solo mediante la referencia a un deseo que se puede explicar la reacción motivacional específica de *S1*. En concreto, parece que semejante reacción sólo podrá explicarse atribuyéndole a *S1* un deseo desmedido de compensar a *S2*.

Partiendo de los casos anteriores, se puede entender cómo es que además de ser compatible con la tesis débil de THM, mi teoría sentimentalista de los juicios morales podría implicar a esa misma tesis, según la cual la posesión de una creencia y un deseo es necesaria para tener motivación. La necesidad de poseer una creencia parece implicarse porque es la posesión de una creencia lo que explica por qué hay casos en los que una emoción y un deseo no son suficiente para generar motivación. Y la necesidad de poseer un deseo parece implicarse porque es la posesión de un deseo lo que permite explicar por qué, dada la posesión de una emoción y una creencia, se adquiere motivación para llevar a cabo un tipo de acción en particular.

6. Lecciones derivadas de las objeciones al racionalismo de Smith

En esta sección, pasaré a mostrar cómo es que mi teoría emotivista de los juicios morales se articula como una mejor solución conciliadora a PM que la teoría racionalista propuesta por Smith. Para ello, mostraré cómo tres objeciones diferentes que afectan a la teoría de Smith no afectan a mi teoría. Primero, una objeción centrada en la identificación de un problema de hiper-intelectualización derivado del tipo de requisitos cognitivos que la teoría de Smith supone que deben de satisfacerse para juzgar moralmente. Después, una objeción que apunta a la incapacidad de la teoría de Smith para dar cuenta de la contingencia inherente al tipo de motivación para la acción que generan los juicios morales. Y por último, una objeción que ataca la viabilidad de que desde la teoría racionalista de Smith se pueda implicar el

internismo moral, en su condición como uno de los enfoques cuya verdad tendría que garantizarse para solucionar PM de forma conciliadora.

6.1. Sentimentalismo y plausibilidad psicológica

En la sección 4.1 del capítulo II, presenté una primera objeción a la teoría racionalista de los juicios morales planteada por Smith como solución al problema moral. A dicha objeción le dí el nombre de "objeción de hiper-intelectualización", pues apunta al reconocimiento de un problema de hiper-intelectualización en la imagen de los juicios que ofrece la teoría racionalista de Smith. Según lo estipulado por Smith, al emitir un juicio moral, un sujeto expresa una creencia acerca de lo que una versión plenamente racional de sí mismo desearía que él hiciera en su versión actual. Semejante concepción parece implicar que, para juzgar moralmente, un sujeto tiene que contar con una gran capacidad de procesamiento cognitivo y manejar conceptos de un alto grado de complejidad. El requisito de procesamiento cognitivo es problemático porque resulta incompatible con el carácter rápido y automático que de hecho manifiestan los juicios morales. Y el requisito de manejo de conceptos complejos es problemático porque resulta razonable pensar que hay casos de sujetos que, pese a no manejar conceptos complejos, pueden emitir juicios morales. Partiendo del desarrollo de esta objeción, llegué a la siguiente lección con respecto a lo que cabría esperar de una buena teoría de los juicios morales:

> ***Plausibilidad psicológica:*** una buena teoría de los juicios morales debería (i) acomodar la aparente rapidez y automaticidad de muchos juicios morales y (ii) no exigir la posesión de conceptos que no estén disponibles para ciertos sujetos capaces de formar juicios morales.

La teoría sentimentalista de los juicios morales que propongo como solución a PM logra incorporar la lección de plausibilidad psicológica. Esto se debe a que al identificar a las emociones morales como bases de los juicios morales, la teoría habilita la posibilidad de entender a los juicios morales como reacciones rápidas y automáticas que no requieren de la posesión de conceptos complejos.

Ahora bien, para comprender por qué mi teoría permite afirmar que los juicios morales no requieren de la posesión de conceptos complejos, recurriré a evidencia

proveniente del ámbito de la psicología del desarrollo. Específicamente, a evidencia descrita por Vaish y Hepach (2020) con respecto al desarrollo de dos emociones morales como la gratitud y la culpa. En lo correspondiente a la gratitud, lo autores indican que las primeras manifestaciones de conducta observable asociadas con dicha emoción moral aparecen alrededor de los 3 años de edad. Por ejemplo, niños de esta edad muestran una tendencia a compartir más con quienes identifican que los han ayudado intencionalmente que con quienes sin intención de ayudar les acercan recursos valiosos. Un poco después, entre los 4 y 5 años de edad, los niños comienzan a valorar positivamente las expresiones de gratitud (más allá de un simple "gracias") de parte de otros sujetos, inclusive más positivamente que las expresiones de felicidad manifestadas en respuesta a acciones que ellos mismos han realizado. Por otra parte, en lo concerniente a la culpa, los autores señalan que las primeras manifestaciones de conducta observable distintivas de esa emoción moral se producen a los 3 años. A esa edad, niños muestran una tendencia, tanto física como verbal, a reparar daños a otros que saben que ellos mismo han causado.[38] Y un poco más adelante, entre los 4 y 5 años, comienza a manifestarse una tendencia a evaluar positivamente las expresiones de culpa que otros hacen.

La evidencia descrita por Vaish y Hepach (2020) respalda la idea de que los juicios morales no requieren de la posesión de conceptos complejos. Esto se debe a que los casos registrados muestran cómo un sujeto puede poseer una emoción moral específica, cuyo contenido incluye ciertos conceptos, incluso cuando dicho sujeto se encuentra en una etapa del desarrollo cognitivo anterior a la que asociamos con la posesión de dichos conceptos. Por ejemplo, el contenido de una emoción como la gratitud involucra el concepto de *reciprocidad*, el cual no resulta plausible que un niño de 4 años posea. Y de manera similar, el contenido de la emoción de culpa involucra el concepto de *reparabilidad* que tampoco es plausible que un niño de 4 años posea. En cualquiera de los casos, el criterio que considero relevante para determinar si un sujeto posee o no un concepto es lo que Evans (1982) denomina "requisito de generalidad". Es decir, la capacidad de recombinar el concepto relevante con otros conceptos.

[38] El reconocimiento de ellos mismos como causantes del daño es el rasgo que permite distinguir a la culpa de otra emoción moral como la simpatía.

6.2. Sentimentalismo y cambio en la justificación

En la sección 4.2 del capítulo II, desarrollé una objeción a la teoría racionalista de los juicios morales de Smith a la que llamé "objeción de justificación contingente". De acuerdo con dicha objeción, dados los términos en los que Smith conceptualiza el fundamento normativo de los juicios morales, su teoría racionalista no puede acomodar la contingencia inherente a la justificación que generan estos juicios. Según lo estipulado por Smith, en la medida en que los juicios morales constituyen razones normativas para la acción, llegar a un juicio moral supone llegar a un veredicto final con respecto a cuáles son los deseos de un agente plenamente racional. De esta manera, un juicio moral genuino es entendido como el producto del ejercicio efectivo del conjunto de capacidad constitutivas de la plena racionalidad. Sin embargo, si los juicios morales se conceptualizan en términos semejantes, entonces, el hecho, constatable en la vida cotidiana, de que el tipo de justificación para la acción que generan estos juicios es susceptible al cambio, no puede dotarse de sentido. Con el objetivo de evitar cometer el mismo error que comete la teoría racionalista de Smith, un buena teoría de los juicios morales debería de atender a la siguiente lección:

> ***Cambio en la justificación:*** una buena teoría de los juicios morales debería de poder acomodar el hecho de que la justificación que un juicio moral puede dar a una acción es susceptible de disminuirse, fortalecerse o cortarse como consecuencia de la interacción con otro juicio moral.

La teoría sentimentalista de los juicios morales que propongo tiene la capacidad de incorporar la lección de "cambio en la justificación". Esto se debe a que permite reconocer que la justificación que genera un juicio moral puede cambiar como consecuencia de la interacción con otro juicio moral. De acuerdo con la teoría, la justificación que ofrecen los juicios morales es un fenómeno derivado de la normatividad prosocial a la que son sensibles las emociones morales en tanto bases de esos juicios. Esa normatividad prosocial se manifiesta en los objetos formales a través de los cuales las emociones morales evalúan a sus objetos intencionales. Así, al evaluar objetos intencionales a través de objetos formales, las emociones morales justifican disposiciones a actuar.

La estrategia que seguiré para mostrar cómo la justificación que genera un juicio moral puede cambiar, se centra en el reconocimiento de la posibilidad de que un mismo sujeto puede realizar diferentes evaluaciones emocionales cuyos respectivos efectos

justificativos interactúan entre sí. La estrategia gira en torno al escenario hipotético de un sujeto *S* que siente culpa por no haber devuelto el dinero que otro sujeto, *P*, le prestó. Dada su emoción de culpa, *S* juzga moralmente incorrecta su acción y encuentra justificación para disculparse con *P*.

Imaginemos ahora una primera variante del caso de *S* en el que, además de sentir culpa por no haberle devuelto el dinero prestado, *S* también siente vergüenza frente a *P* porque es consciente de que él tiene más dinero que *P*. En este escenario, puesto que *S* no solo siente culpa respecto de su acción hacia *P*, sino que además siente vergüenza frente a *P*, la justificación que *S* tiene para disculparse con *P* es mayor que la que tenía en el caso original, cuando únicamente sentía culpa. En resumen: dos juicios morales de *S*, basados en dos emociones morales diferentes, interactúan entre sí y en consecuencia su justificación inicial se ve fortalecida.

Ahora bien, en una segunda variante del caso de *S*, supongamos que, si bien *S* siente culpa por no haberle devuelto a *P* el dinero que le prestó, *S* también siente indignación hacia *P* como consecuencia del hecho de que *P* acostumbra burlarse de la estatura de *S*. En este escenario, aunque la culpa que siente por no haber regresado el dinero justifica el que *S* se disculpe con *P*, la indignación que siente hacia *P* justifica el que evite interactuar con él. En resumen: dos juicios morales de *S*, basados en dos emociones morales diferentes, interactúan entre sí y en consecuencia su justificación inicial se ve disminuida.

Finalmente, imaginemos una tercera variante del caso de *S* en el que, si bien *S* sentía culpa por no haberle devuelto a *P* el dinero que le prestó, *S* ha comenzado a sentir ira hacia *P* después de enterarse de que durante años, *P* le ha estado robando. En este escenario, la justificación que *S* tenía para actuar en atención a la culpa que sentía originalmente hacia *P*, queda neutralizada como consecuencia de que la ira que siente hacia *P* lo justifica a buscar resarcir el daño que *P* le ha causado. En resumen: dos juicios morales de *S*, basados en dos emociones morales diferentes, interactúan entre sí y en consecuencia su justificación inicial se corta.

En su conjunto, las tres variantes descritas del escenario hipotético protagonizado por *S* permiten mostrar cómo la teoría sentimentalista que defiendo tiene la capacidad de incorporar la lección de "cambio en la justificación". Más específicamente, partiendo del reconocimiento de la posibilidad de que un mismo sujeto puede ser sensible a diferentes emociones morales, la teoría consigue acomodar el hecho de que la justificación que genera un juicio moral puede fortalecerse, disminuirse e incluso cortarse.

6.3. Sentimentalismo y modo de presentación inadecuado

En la sección 4.3 del capítulo II, desarrollé una objeción a la teoría racionalista de los juicios morales de Smith a la que llamé "objeción de modo de presentación inadecuado". A grandes rasgos, lo que dicha objeción establece es que el contenido de las creencias expresadas por juicios morales utiliza un modo de presentación inadecuado. En particular, un modo de presentación inadecuado en el sentido de que impide que la teoría de Smith pueda dar implicar al internismo moral. En última instancia, esto resulta problemático pues solo si la teoría es capaz de implicar el internismo moral, podrá funcionar como una solución conciliadora a PM. Es decir, como una solución que rescata la verdad de las tres proposiciones que conforman el problema.

La objeción de modo de presentación inadecuado derivó en la siguiente lección con respecto a lo que cabría esperar de una teoría de los juicios morales:

> ***Modo de presentación actual:*** una buena teoría de los juicios morales debería de concebir el contenido de esos juicios bajo un modo de presentación que, o bien describa la manera como se concibe actualmente el sujeto que juzga, o bien esté conectado con lo que es actualmente, sin introducir subrepticiamente elementos motivadores no explicados.

La teoría sentimentalista que propongo constituye una teoría capaz de incorporar la lección anterior. Esto se debe a que al identificar a las emociones como bases de los juicios morales, la teoría asegura una conexión entre el juicio moral de un sujeto y la manera en que ese sujeto se concibe a sí mismo en la actualidad. El factor clave en la concreción de esta conexión es la condición de las emociones como estados que reflejan las preocupaciones de quien las posee, algo en lo cual profundicé en la sección 2.5 de este capítulo. Recordando lo dicho en esa sección, las emociones de un sujeto reflejan sus preocupaciones en el sentido de que son sus preocupaciones las que, al menos en parte, determinan cuales son las emociones que experimenta. La relevancia de esta fundamentación de las emociones en preocupaciones se hace patente una vez que se toma en cuenta que las preocupaciones de un sujeto reflejan qué tipo de persona es ese sujeto en la actualidad. Para entender este último punto hay que comprender cómo se producen las preocupaciones.

De acuerdo con (Roberts, 2003, p.142), las preocupaciones de un sujeto pueden tener dos orígenes diferentes. Por un lado, pueden originarse como consecuencia de factores

biológicos en el sujeto. Y por otro lado, pueden surgir en atención a factores aprendidos por el sujeto. De una u otra manera, una preocupación refleja aquello que es importante en su actualidad. Esta afirmación puede entenderse en términos representacionales. De esta manera, lo que se afirma es que una preocupación específica de un sujeto refleja lo que, según una representación del propio sujeto, es importante para él. Por ejemplo, la preocupación de *S* por comer algo, refleja una representación que *S* tiene de sí mismo como alguien hambriento. Y de la misma manera, la preocupación de *S* por que lo que vaya a comer sea algo sano, refleja una representación que *S* tiene de sí mismo como alguien que cuida su salud. Sin embargo, la manera en que mi teoría entiende la capacidad de las preocupaciones para reflejar lo que es importante para un sujeto en su actualidad es en términos no-representacionales. De acuerdo con esta manera de entender la cuestión, una preocupación específica de un sujeto *S* refleja lo que es importante para *S* en su actualidad, simplemente porque esa preocupación está constituida por lo que *S* es en la actualmente. Así, se afirma que la preocupación de *S* refleja lo que es importante para *S* independientemente de cómo *S* se represente a sí mismo.

En lo particular de preocupaciones que fundamentan emociones morales, la cuestión se articula de la misma manera. Por ejemplo, mi preocupación por el bienestar de mi madre, que se refleja en la emoción moral de gratitud, refleja el hecho de que en la actualidad, soy una persona que aprecia el apoyo que recibe de otras personas. Y de manera similar, mi preocupación por no consumir lujos innecesarios que se refleja en la emoción moral de vergüenza, refleja a su vez el hecho de que en la actualidad, soy una persona que se reconoce como privilegiada viviendo en una sociedad profundamente desigual. En general, eso que uno es actualmente que se refleja en el tipo de preocupaciones que fundamentan emociones morales es algo que determina un proceso de entrenamiento moral. Por ejemplo, el hecho de que actualmente soy una persona que aprecia cuando recibe apoyo de otras personas, proviene en última instancia de la experiencia de ver a mi madre cuidar a su propia madre. Y el hecho de que soy una persona que se reconoce como alguien privilegiado viviendo en una sociedad profundamente desigual, es consecuencia del contacto con relatos sobre la vida cotidiana de otras personas que refieren a dificultades con las que yo no me encuentro.

Partiendo de la descripción anterior acerca del sentido en el cual las emociones morales, en tanto fundamentadas en preocupaciones, reflejan lo que es importante para un sujeto en su actualidad, procedo a especificar cómo mi teoría sentimentalista incorpora la lección de "modo de presentación actual". En concreto, al afirmar que los juicios morales se explican por la posesión de emociones morales, la teoría asegura que haya un sentido en el cual el

contenido de esos juicios sostenga una conexión con lo que la persona que los emite es actualmente.

7. Conclusiones

El objetivo de este capítulo ha sido desarrollar una teoría de los juicios morales que pueda funcionar como una solución conciliadora al problema moral mejor que la propuesta por Smith (1994). La teoría desarrollada se configura como una forma de sentimentalismo moral. Su tesis principal es que si una determinada actitud es un juicio moral, entonces dicha actitud tiene una (o varias) base(s) en una emoción moral. Según lo mostrado en el capítulo, la teoría constituye una solución conciliadora al problema moral en el sentido de que implica al cognitivismo moral, al internismo moral y la teoría humeana de la motivación (en su versión débil). Pero además, mejora a la teoría racionalista de Smith, en tanto no enfrenta tres objeciones que sí afectan a esa teoría.

CONCLUSIONES

Este trabajo ha estado dedicado a explorar la posibilidad de ofrecer una respuesta satisfactoria al problema moral según lo describe Smith (1994). De acuerdo con dicho problema, si aceptamos la teoría humeana de la motivación, entonces, contrariamente a lo que la práctica moral ordinaria invita a pensar, no podemos atribuir a los juicios morales tanto una función descriptiva como una función práctica. A lo largo del trabajo he considerado a profundidad dos propuestas de solución diferentes. Primero, una propuesta consistente en una teoría racionalista de los juicios morales. Esta propuesta, desarrollada por el mismo Smith, afirma que para resolver el problema moral deberíamos pensar que los juicios morales expresan un cierto tipo de creencias, específicamente, creencias acerca de los deseos de un agente plenamente racional. Pese a que consigue acomodar la verdad de las tres tesis que conforman el problema moral, esta propuesta racionalista presenta deficiencias explicativas considerables que demeritan su plausibilidad. Por otra parte, la segunda propuesta de solución al problema moral que he abordado es una propuesta original que se articula como una teoría sentimentalista de los juicios morales. Desde la perspectiva de esta teoría, los juicios morales se explican por la posesión de emociones morales. Al igual que la teoría racionalista de Smith, esta teoría sentimentalista puede acomodar la verdad de las tres tesis que conforman el problema moral. Sin embargo, en contraste con la teoría de Smith, esta teoría no presenta las deficiencias explicativas que sí afectan al enfoque racionalista.

Más allá de lo que ha logrado mostrar, el trabajo deja abiertas algunas cuestiones que de ser abordadas podrían enriquecer la comprensión teórica de los juicios morales. Una primera cuestión que ha quedado abierta concierne a la especificación del tipo de teoría de las emociones que de mejor manera podría sustentar una teoría sentimentalista de los juicios morales como la aquí defendida. Para el desarrollo de este trabajo, he optado por formular una concepción general de las emociones. Esta concepción general caracteriza a los estados emocionales a través de la identificación de componentes que, desde diversas teorías, comúnmente se les atribuyen. Asimismo, para captar el sentido en el cual las representaciones propias de las emociones pueden ser de relevancia moral, he recurrido al concepto de emociones morales de Haidt (2003). En su conjunto, la concepción general ofrecida y la referencia a la categoría más específica de las emociones morales, han permitido articular una imagen de las emociones lo suficientemente robusta para cumplir con los objetivos específicos de esta investigación. Sin embargo, para aumentar el potencial explicativo de mi teoría, pienso que sería importante ofrecer una historia completa del

funcionamiento de cada uno de los componentes de los estados emocionales. Es decir, sería importante introducirse en el debate en filosofía de la mente concerniente a la naturaleza de las emociones.

Por otro lado, una segunda cuestión que ha quedado abierta es la relativa a la pertinencia de la distinción entre tipos de estados mentales en términos de dirección de ajuste. Dicha distinción, fundada en ciertas intuiciones teóricas plausibles pero también cuestionables, cumple un rol fundamental en la configuración del problema moral. Pues es con base en ella que se da sustento a la afirmación de que un mismo estado mental no puede desempeñar tanto una función descriptiva como una función práctica. Partiendo del reconocimiento de su protagonismo en el planteamiento de la cuestión, considero que una estrategia prometedora para resolver la tensión capturada por el problema moral consistiría en atacar esa distinción de dirección de ajuste entre estados mentales.

BIBLIOGRAFÍA

Anscombe, G. E. M. (1991). *Intención* (A. I. Stellino, trad.). Paidós. (Obra original publicada en 1957).

Ayer, A. J. (1954). Freedom and Necessity. En *Philosophical Essays*. MacMillan.

Blackburn, S. (1984). *Spreading the Word: Groundings in the Philosophy of Language*. Oxford University Press.

Brady, M. S. (2024). An Overview of Contemporary Theories of Emotions in Philosophy. En A. Scarantino (Ed.), *Emotion Theory: The Routledge comprehensive guide* (pp. 215–231). Routledge.

Brink, D. O. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge University Press.

Dancy, J. (2004). *Ethics Without Principles*. Oxford University Press.

Davidson, D. (1980). Intending. En D. Davidson, *Essays on Actions and Events* (pp. 83–102). Oxford University Press. (Obra original publicada en 1978)

Deonna, J. A., y Teroni, F. (2012). *The Emotions: A Philosophical Introduction*. Routledge.

Echeverri, S. (2019). Emotional justification. *Philosophy and Phenomenological Research, 98*, 541–566. https://doi.org/10.1111/phpr.12453

Evans, G. (1982). *The Varieties of Reference*. Clarendon Press

Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68(1), 5–20.

Frege, G. (1996). Sobre el sentido y la referencia. En L. M. Valdés Villanueva (Ed.), *La búsqueda del significado* (pp. 23-45). Tecnos. (Obra original publicada en 1892).

Gauthier, D. (1975). Reason and Maximization. Canadian Journal of Philosophy 4(3), 411-433.

Geach, P. T. (1960). Ascriptivism. *The Philosophical Review, 69*(2), 221–225.

Geach, P. T. (1965). Assertion. *The Philosophical Review, 74*(4), 449–465.

Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.

Gibbard, A. (2003). *Thinking How to Live*. Harvard University Press.

Goldie, P. (2000). *The Emotions: A Philosophical Exploration*. Oxford University Press. https://doi.org/10.1093/0199253048.001.0001

Haidt, J. (2003). The Moral Emotions. En R. J. Davidson, K. R. Scherer, y H. H. Goldsmith (Eds.), *Handbook of Affective Sciences* (pp. 852–870). Oxford University Press.

Haidt, J. (2019). *La mente de los justos: Por qué la política y la religión dividen a la gente sensata* (A. García Maldonado, Trad.). Ediciones Deusto. (Obra original publicada en 2012).

Hume, D. (1978). *A Treatise of Human Nature* (L. A. Selby-Bigge, Ed.; 2.ª ed.). Oxford University Press. (Trabajo original publicado en 1739).

Johnston, M. (1989). Dispositional Theories of Value. *Proceedings of the Aristotelian Society, Supplementary Volumes, 63*, 139–174.

Korsgaard, C. M. (1986). Skepticism About Practical Reason. *The Journal of Philosophy, 83*(1), 5–25.

Kriegel, U. (2012). Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction. *Australasian Journal of Philosophy, 90*(3), 469–486.

Lewis, D. (1989). Dispositional Theories of Value. *Aristotelian Society Supplementary Volume*, 63(1), 113–139.

Nagel, T. (1970). *The Possibility of Altruism*. Oxford University Press.

Neu, J. (2000). *A Tear is An Intellectual Thing*. Oxford University Press.

Nussbaum, M. (2001). *Upheavals of Thought*. Cambridge University Press.

Pacherie, E. (2002). The Role of Emotions in the Explanation of Action. *European Review of Philosophy, 5*, 53–92.

Pettit, P., y Smith, M. (1993). Practical Unreason. *Mind, 102*(405), 53–79.

Platts, M. (1979). *Ways of Meaning*. Routledge & Kegan Paul.

Platts, M. (1981). Moral Reality and the End of Desire. En M. Platts (Ed.), *Reference, Truth and Reality: Essays on the Philosophy of Language* (pp. 69–82). Routledge & Kegan Paul.

Prinz, J. (2006). The Emotional Basis of Moral Judgments. *Philosophical Explorations, 9*(1), 29–43.

Prinz, J. (2007). *The Emotional Construction of Morals*. Oxford University Press.

Rawls, J. (1951) Outline of a Decision Procedure for Ethics. *The Philosophical Review*, 60(2), 177–197.

Roberts, R. (2003). *Emotions: An essay in aid of moral psychology*. Cambridge University Press.

Sayre-McCord, G. (1997). The Metaethical Problem. *Ethics, 108*(1), 55–83.

Scarantino, A. (2014). The Motivational Theory of Emotions. En J. D'Arms y D. Jacobson (Eds.), *Moral Psychology and Human Agency* (pp. 156–185). Oxford University Press.

Scarantino, A., y De Sousa, R. (2021). Emotion. En E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (edición de verano de 2021). https://plato.stanford.edu/archives/sum2021/entries/emotion/

Schroeder, M. (2007). What is the Frege-Geach problem? *Philosophy Compass, 3*(4), 703–720.

Smith, A. (2002). *The Theory of Moral Sentiments*. Cambridge University Press. (Obra original publicada en 1759).

Smith, M. (1994). *The Moral Problem*. Blackwell.

Stocker, M. (1979). Desiring the Bad: An Essay in Moral Psychology. *The Journal of Philosophy*, *76*(12), 738–753.

Stroud, B. (1977) *Hume*. Routledge and Keagan Paul.

Tappolet, C. (2016). *Emotions, Value, and Agency*. Oxford University Press.

Tappolet, C. (2023). *Philosophy of Emotion: A Contemporary Introduction*. Routledge.

Vaish, A., & Hepach, R. (2020). The development of prosocial emotions. *Emotion Review*, *12*(4), 259–273. https://doi.org/10.1177/1754073919885014

Watson, G. (1975) Free Agency. *The Journal of Philosophy*, *72*(8), 205–220. https://doi.org/10.2307/2024703

Williams, B. (1980). Internal and External Reasons. En *Moral Luck.* Cambridge University Press; 1981. pp. 101–113.

Woods, J. (2017). The Frege-Geach problem. En T. McPherson y D. Plunkett (Eds.), *The Routledge handbook of metaethics* (pp. 226–242). Routledge.