



UNIVERSIDAD NACIONAL DE COLOMBIA

# **El Problema de la Semántica del Lenguaje Moral en la Discusión Metaética Contemporánea**

**Sergio Almeida Moreno**

Universidad Nacional de Colombia

Departamento de Filosofía

Bogotá D.C., Colombia

2012

# **El Problema de la Semántica del Lenguaje Moral en la Discusión Metaética Contemporánea**

**Sergio Almeida Moreno**

Tesis presentada como requisito parcial para optar al título de:

**Magister en Filosofía**

Director:

Ph.D. Juan José Botero Cadavid

Línea de Investigación:

Metaética

Universidad Nacional de Colombia

Departamento de Filosofía

Bogotá D.C., Colombia

2012

## **Agradecimientos**

Quiero agradecer a la División de Investigación de la Universidad Nacional, sede Bogotá, por el apoyo financiero recibido para la realización de esta tesis, a través de la Convocatoria de Apoyo a Tesis de Posgrado 2009, Proyecto DIB 8005127.

## Resumen

El propósito de la tesis es exponer de manera clara el problema que hay en torno a la especificación de la semántica del lenguaje moral y proponer una teoría capaz de solucionar dicho problema, inspirada en el intuicionismo de G.E. Moore. En el capítulo 1 examino la formulación de Moore del Intuicionismo y las críticas que se han hecho al mismo. A partir de ello, planteo el que a mi modo de ver es el problema central en torno a la semántica del lenguaje moral. En el capítulo 2 doy una breve caracterización de las teorías Antirrealistas Subjetivistas, las cuales inicialmente se presentan como la mejor alternativa para superar los problemas del Intuicionismo. A continuación, presento la crítica que John McDowell ha hecho a las mismas y muestro la necesidad de buscar un modelo Antirrealista que no caiga en el subjetivismo. En el capítulo 3 comienzo examinando los requisitos que, según parece, debe satisfacer el lenguaje moral para ser objetivo. A continuación, propongo un modelo para explicar la semántica moral, el cual se basa en los planteamientos de Michael Dummett acerca del Antirrealismo. Finalmente, en el capítulo 4 presento la discusión en torno al Internalismo, indicando los argumentos que permiten defenderlo. Este capítulo es importante porque, como se puede ver desde el capítulo 1, considero que cualquier teoría que pretenda explicar satisfactoriamente el lenguaje moral necesariamente debe dar cuenta de dicha tesis.

**Palabras clave:** Metaética, Intuicionismo, Expresivismo, Antirrealismo Moral, Intersubjetividad, Internalismo

## Abstract

The aim of this work is to expose in a clearly way the problem that surrounds the specification of the moral language' semantics and propose a theory able to solve the problem that is inspired in G.E. Moore's intuitionism. In chapter 1 I explore Moore's formulation of intuitionism and the critics that It has received. On that basis, I'll present which, under mi point of view, is the central problem of moral language' semantics. In chapter 2 I offer a brief characterization of Antirealist Subjectivist theories, which are originally presented as the best alternative solution to the intuitionism problems. Below, I present the critics that John McDowell has done to that theories and I show the need to find an antirealist model which don't implicate subjectivism. At chapter 3 I examine the requirements that, it appears, must satisfies moral language to be considered objective. Then, I'll propose a model able to explain moral semantics, that is based on Michel Dummett's formulation of Antirealism. Finally, In chapter 4 I present the discussion about Internalism, indicating the arguments that let it to be defended. The importance of these chapter reside in, how you can infer from chapter 1, I consider that any theory that pretends to explain adequately moral language must necessarily give an account of that thesis.

**Keywords:** Meta-ethics, Intuitionism, Expressivism, Moral Antirealism, Intersubjectivity, Internalism

# Tabla de Contenidos

|  |    |
|--|----|
| Introducción.....  | 1  |
| 1. El Intuicionismo.....   | 4  |
| 1.1. La Falacia Naturalista y el Intuicionismo de Moore.....         | 4  |
| 1.2. El Argumento de la Pregunta Abierta y el Internalismo.....      | 11 |
| 1.3. Epistemología y Metafísica Especiales.....                      | 14 |
| 1.4. El Problema Moral.....  | 17 |
| 2. Los Inconvenientes del Antirrealismo Subjetivista.....            | 20 |
| 2.1. Los Postulados Básicos del Expresivismo.....                    | 20 |
| 2.2. La Crítica de McDowell al Expresivismo.....                     | 22 |
| 3. Objetividad, Antirrealismo y Experiencia Intersubjetiva.....      | 29 |
| 3.1. El Problema de la Objetividad del Lenguaje Moral.....           | 29 |
| 3.2. Criterios de Objetividad.....                                   | 32 |
| 3.3. El Modelo Antirrealista de Dummett.....                         | 34 |
| 3.4. La Experiencia de Intersubjetividad.....                        | 39 |
| 4. La Polémica Internalismo vs Externalismo.....                     | 44 |
| 4.1. Dos Versiones del Internalismo.....                             | 44 |
| 4.2. La Defensa de Hare del Internalismo.....                        | 48 |
| 4.3. El Amoralismo de D.O. Brink y el Externalismo.....              | 49 |
| 4.4. La Defensa de Smith del Internalismo y el Fetichismo Moral..... | 52 |
| Bibliografía.....  | 54 |

# INTRODUCCIÓN

El propósito de este trabajo es esbozar una solución al problema de cómo especificar la semántica del lenguaje moral. Dicho problema suele presentarse en términos de la dicotomía Cognitivismo – No- Cognitivismo. Así, por ejemplo, en su libro *An Introduction to Contemporary Metaethics*, Alexander Miller parte de una clasificación de las diversas teorías que han tratado de responder a esta cuestión, para presentar los principales interrogantes que hay en torno a la semántica del lenguaje moral.

La primera gran división, correspondiente a las teorías *Cognitivistas* y *No-Cognitivistas*, se da en virtud de la pregunta ¿Los juicios morales poseen condiciones de verdad? O en otras palabras, ¿Los juicios morales son aptos para la verdad (truth-apt)? Naturalmente, si se responde afirmativamente, se estará afirmando que los juicios morales expresan creencias, puesto que las creencias son estados mentales que típicamente poseen condiciones de verdad. Entonces, la división también se puede trazar tomando en consideración la pregunta ¿Los juicios morales expresan creencias? Las teorías *Cognitivistas* son aquellas que responden afirmativamente a tales preguntas. Ellas se definen por atribuirle condiciones de verdad a los juicios morales y por considerarlos como expresiones de creencias. En contraste, las teorías *No-Cognitivistas* consideran que los juicios morales no tienen condiciones de verdad. En vez de expresar creencias, los juicios morales expresan estados mentales que no pueden ser evaluables en términos de verdad o falsedad (a diferencia de las creencias que pueden hacerlo en relación con el mundo) Algunos ejemplos de estos estados serían deseos, sentimientos y emociones, a los que se les reconoce como estados mentales no-cognitivos. Entre las teorías No-Cognitivistas se encuentran: el Emotivismo de Ayer y Stevenson, el Prescriptivismo de R.M. Hare (1952) y el Expresivismo Normativo de Allan Gibbard (1990).

Según Miller, considerar otros interrogantes que se le pueden plantear a la semántica del lenguaje moral puede conducirnos a trazar algunas divisiones al interior de los grandes grupos mencionado. Por ejemplo la cuestión de si las condiciones de verdad de un juicio moral se refieren a *hechos en el mundo* o a *estados u opiniones de las personas* que emiten esos juicios puede llevarnos a dividir a las teorías *Cognitivistas* entre *Cognitivismo Fuerte* y *Cognitivismo Débil*.

Quienes responden que las condiciones de verdad de los juicios morales los hacen responsables frente a una realidad independiente de la mente humana sostienen una forma de *Cognitivismo Fuerte*. Las variedades de Cognitivismo fuerte son numerosas y la taxonomía continuara tomando en consideración distinciones que se pueden hacer al interior de las teorías de este tipo. Por su parte, quienes consideran que las condiciones de verdad de los juicios morales están inevitablemente ligadas a las opiniones de las personas que los sostienen, i.e., que los juicios morales no son responsables frente a una realidad independiente sino frente a hechos constitutivos de la psicología humana, sostienen una forma de *Cognitivismo Débil*. Esta posición ha sido asociada al Anti-Realismo sostenido por Crispin Wright (1992).

La otra pregunta adicional que se formulan los cognitivistas fuertes, y que permite elaborar una nueva distinción al interior del grupo, es si existen o no juicios morales

verdaderos. Quienes responden afirmativamente a esta cuestión consideran que en el mundo *hay hechos morales* que, efectivamente, hacen verdaderos a algunos de nuestros juicios morales. Por tal razón, quienes sostienen este tipo de teorías son realistas. La mayor parte de variedades de Cognitivismos Fuertes es de este corte. Al interior de estas teorías aún es posible hacer divisiones más precisas.

No obstante, John Mackie (1977) ha sostenido que, aunque el Cognitivismos Fuerte está en lo correcto al especificar las condiciones de verdad de los juicios morales de tal modo que los hace responsables frente a una realidad independiente, es erróneo suponer que hay juicios morales verdaderos. Mackie sostiene que todos los juicios morales son falsos, puesto que no existen propiedades morales en el mundo. Según él, en primer lugar, pensar que tales propiedades existen implica sobrecargar la metafísica y la epistemología: tendríamos que admitir que en el mundo hay propiedades que tienen características especiales, diferentes de las de las propiedades naturales (i.e., propiedades que guía directamente nuestras acciones), y además, tendríamos que admitir que poseemos algún tipo de capacidad cognitiva especial que nos da acceso a epistémico a ellas. Y en segundo lugar, si tales propiedades existieran, no habría desacuerdos morales. En vista de que ambas cosas parecen absurdas, Mackie concluye que las propiedades morales no pueden existir. En consecuencia que todos los juicios morales son falsos. Esta posición es conocida como la *Teoría del Error*.

Todos los Cognitivistas Fuertes Realistas están de acuerdo con la tesis de que los juicios morales poseen condiciones de verdad y que hay hechos morales, pertenecientes a la estructura del mundo, que los hacen verdaderos. Las diferencias entre los Cognitivistas de este corte, entonces, se dan en virtud de cómo ellos conciben la naturaleza de tales hechos morales.

La primera división entre Cognitivistas Fuertes Realistas tiene que ver con la pregunta ¿Las propiedades morales son propiedades naturales o propiedades *sui generis*? Quienes consideran que las propiedades morales son propiedades *sui generis* son *Cognitivistas No-Naturalistas*. Entre estas teorías se encuentran el Intuicionismo de Moore y la Teoría Sensibilista de John McDowell (1998). Moore consideraba que las propiedades morales era especiales en el sentido en que no podían establecer relaciones causales con otras propiedades en el mundo. Sin embargo, él nunca explicó plenamente cuál era, a su forma de entender, la naturaleza de tales propiedades. McDowell, por su parte, ha sostenido en numerosos ensayos, que las propiedades morales son propiedades secundarias, análogas a propiedades como el color. Según McDowell, ellas son propiedades disposicionales, esto es, propiedades que se fundamentan en propiedades objetivas de la realidad, pero que se definen por el hecho de que causan ciertas experiencias en nosotros. En contraste, quienes consideran que las propiedades morales son naturales son *Cognitivistas Naturalistas*. Hay, sin embargo, dos formas de Naturalismo, correspondientes a dos formas diferentes de concebir la relación entre propiedades naturales y propiedades morales. Algunos autores como Peter Railton (1986) consideran que las propiedades morales son *reducibles* a propiedades naturales. Esto implica que es posible dar una definición naturalista del tipo que Moore rechaza mediante el APA. Este tipo de teorías se conoce como *Naturalismo Reduccionista*. Por otra parte, hay autores como Nicholas Sturgeon (1988), Frank Jackson y Philip Pettit (2004) que piensan que las propiedades morales son propiedades naturales de un tipo único, no reducibles a otras propiedades morales. La tesis central de este tipo de Naturalismo es que las propiedades morales tienen la capacidad de sostener relaciones causales y que son relevantes para la explicación de nuestra experiencia moral. Este tipo de Cognitivismos se conoce como *Realismo Cornell*, puesto que tuvo su origen en la Universidad de Cornell.



En vista de esta pluralidad de teorías, la cuestión con respecto a la semántica del lenguaje moral puede plantearse simplemente diciendo: ¿cuál de estas teorías es la correcta, aquella que da la explicación más satisfactoria del lenguaje moral? Sin embargo, creo que al plantear la cuestión de este modo, a partir una taxonomía hecha sobre la base de la dicotomía Cognitivismo – No-Cognitivismo, no se aprecian con claridad las diversas cuestiones que hacen de la tarea de especificar la semántica moral algo tan problemático. Por tal razón, el primer objetivo de este trabajo es lograr una formulación del problema que deje ver todo lo que hay en juego aquí. El primer capítulo está dedicado a cumplir con esta tarea. Adicionalmente, como hemos visto la taxonomía de las teorías que hay en torno a esta cuestión es demasiado amplia. En ese sentido, es muy probable que un examen detallado y una discusión minuciosa de las mismas haga este trabajo algo tan complejo, que nos lleve a perdernos en los detalles y a perder de vista los aspectos esenciales de la discusión. Teniendo en cuenta esto, un segundo propósito del trabajo es, prescindiendo un poco del trabajo de interpretar y exponer a todos estos autores, esbozar una propuesta que, a mi modo de ver, puede solucionar en gran medida el problema. Considero que el Intuicionismo es la posición más acertada y que la propuesta que presentaré a continuación se aproxima bastante a él. por tal razón, creo que los planteamientos que se encuentran acá pueden ser vistos como una defensa, muy modesta, del Intuicionismo.

En el capítulo 1 examinaremos la formulación de Moore del Intuicionismo y las críticas que le han formulado al mismo. A partir de ello plantearemos el que, a mi modo de ver, es el problema central en torno a la semántica del lenguaje moral. En el capítulo 2 daremos una breve caracterización de las teorías Antirrealistas Subjetivistas, las cuales inicialmente se presentan como la mejor alternativa para superar los problemas del Intuicionismo. A continuación, presentaremos la crítica que John McDowell ha hecho a las mismas y evidenciaremos la necesidad de buscar un modelo Antirrealista que no caiga en el subjetivismo. En el capítulo 3 comenzaremos examinando los requisitos que, según parece, debe satisfacer el lenguaje moral para ser objetivo. A continuación, propondremos un modelo para explicar la semántica moral, el cual se basa en los planteamientos de Michael Dummett acerca del Antirrealismo. Finalmente, en el capítulo 4 presentaremos la discusión en torno al Internalismo, mostrando los argumentos que permiten defenderlo. Este capítulo es importante por qué, como se verá desde el capítulo 1, considero que cualquier teoría que pretenda explicar satisfactoriamente el lenguaje moral necesariamente debe dar cuenta de dicha tesis.

# 1. El Intuicionismo

En la Introducción dijimos que Moore fue el primer filósofo contemporáneo en resaltar la importancia de las investigaciones acerca de la semántica del lenguaje moral, con miras a la resolución de los problemas propios de la ética práctica. En ese sentido, Moore fue el primero en proponer una teoría metaética acerca del lenguaje moral. Dicha teoría es, precisamente, el Intuicionismo. En gran parte, las teorías metaéticas posteriores, como el Expresivismo (en sus diferentes formas) y la Teorías del Error, han surgido como respuesta a las dificultades que se han encontrado en la formulación de Moore del Intuicionismo. Por tal razón, una buena estrategia para adentrarse en la discusión metaética contemporánea, es comenzar planteando el problema que hay en torno a la semántica del lenguaje moral, a partir de las dificultades que se han encontrado en el Intuicionismo. El propósito de este capítulo es, precisamente, hacer una aproximación a esta teoría, que nos permita introducir el problema del que se han ocupado los teóricos en la metaética contemporánea.

## 1.1. La Falacia Naturalista y el Intuicionismo de Moore

Buena parte de los *Principia Ethica* está dedicada a revelar la naturaleza y las consecuencias de un error conceptual que, según Moore, está a la base de todas las teorías éticas Naturalistas. Tal error es lo que él denomina “Falacia Naturalista”. A partir de estas consideraciones, Moore propone al Intuicionismo, como una doctrina metaética que no incurre en la Falacia Naturalista y que, por consiguiente, no debe asumir las consecuencias negativas derivadas de ella. Teniendo en cuenta esto, es claro que una aproximación al Intuicionismo debe comenzar por considerar en qué consiste la Falacia Naturalista. Esta sección está dedicada a examinarla.

A fin de exponer la Falacia Naturalista, la primera cuestión que debemos abordar es: ¿qué es el Naturalismo? o, en otras palabras, ¿qué afirmación o afirmaciones constituyen la doctrina conocida como Naturalismo? Consideremos algunas de las teorías a las que tradicionalmente se les da el rótulo de Naturalistas. Según Moore, entre ellas, las más representativas son las Teorías del Interés y el Utilitarismo. En las Teorías del Interés, la palabra «bueno» se define en términos de estados psicológicos como deseos o actitudes de aprobación. Por ejemplo, según Hobbes, «bueno» significa «deseado por mí»; mientras que, según Hume, «bueno» significa «aprobado por la mayor parte de la gente» (Cf. Stevenson 1937, p. 270). Por su parte, en el Utilitarismo, la palabra «bueno» se define en términos de placer y dolor. Esto puede verse en las obras de Bentham y Mill, en las que se dice que «bueno» significa «que produce placer» (o «que es placentero»). Estos ejemplos muestran un elemento común a las dos clases de teorías, propio del Naturalismo, a saber: la idea de que es posible *definir* la palabra «bueno». Esta es, a los ojos de Moore, la afirmación fundamental del Naturalismo. Sin embargo, aún debemos preguntarnos ¿qué es exactamente lo que los Naturalistas quieren decir, cuando afirman que es posible definir la palabra «bueno»? o, de otro modo, ¿cuál es la noción de «definición» que está presente en sus consideraciones? Al respecto, Moore piensa que es

importante hacer la siguiente aclaración: no debe pensarse que el propósito de los naturalistas es ofrecer una definición léxica de la palabra «bueno». Una definición de este tipo no especifica propiamente el *significado* de una palabra; sino que simplemente, describe el modo en que ella es *usada* por los hablantes de una lengua. Por ello, tales definiciones aparecen en los diccionarios y su formulación se hace a la luz de ciertas investigaciones de carácter empírico, llevadas a cabo por los lexicógrafos. Con estas afirmaciones, Moore deja en claro que, para él, el uso de una palabra no determina su significado; por lo cual, las consideraciones acerca del uso, así como las definiciones léxicas, no tienen ningún valor para la semántica filosófica. En contraste, cuando los Naturalistas dicen que es posible definir la palabra «bueno», piensan que es posible formular una expresión, tal que describa *la naturaleza del objeto o idea* para cuya representación se utiliza generalmente la palabra «bueno» (Cf. Moore 1903, §6, pp. 28-9). Según esto, de acuerdo con el Naturalismo, la expresión empleada para definir «bueno» debe referirse al mismo objeto al que se refiere esta última, aunque lo describa de un modo más preciso; en otras palabras, tal expresión debe ser *sinónima* de «bueno». En consecuencia, las definiciones propuestas por las teorías Naturalistas deben ser vistas como especificaciones del *significado* de la palabra «bueno»; de modo que, el Naturalismo debe concebirse como una doctrina metaética, cuyo propósito principal es explicar la semántica del lenguaje moral.

Una segunda afirmación, constitutiva del Naturalismo, es que existen enunciados éticos *analíticos*. Esta afirmación se sigue inmediatamente de la anterior. En efecto, si es cierto que en la definición que propone todo Naturalista el *definiendum* y el *definiens* son expresiones sinónimas, entonces el enunciado que expresa tal definición siempre será analítico. Los Naturalistas, en consecuencia, deben sostener que existe, al menos, un enunciado ético analítico, a saber, su propia definición de «bueno».

Una tercera afirmación, propia del Naturalismo, es que «bueno» es un *término complejo*. Evidentemente, para entender por qué el Naturalismo debe sostener esta idea, es necesario considerar la distinción que hace Moore entre términos simples, por una parte, y términos complejos, por la otra. Según Moore, los términos simples son términos que se refieren a *objetos simples de la experiencia*, esto es, a objetos “[...] en los que uno puede pensar o percibir, pero de tal naturaleza que nunca pueden transmitirse a alguien incapaz de pensar en ellos o de percibirlos.” (*Idem* §7, p. 29). Así, por ejemplo, «amarillo» es un término simple; porque se refiere a un objeto simple de la experiencia, a saber, el color amarillo, el cual no puede ser conocido por alguien que desde siempre haya carecido del sentido de la vista. En contraste, los términos complejos son términos que se refieren a objetos compuestos por partes, esto es, a *objetos que pueden describirse*, mediante una relación de las partes que los componen, y de un modo tal que alguien que nunca haya tenido experiencia de ellos pueda entender qué clase de objetos son (*Idem* §7, p. 30). Según Moore, la descripción de un objeto complejo debe formularse, en última instancia, empleando términos simples, ya que “[éstos] constituyen los últimos términos con los que ha de hacerse referencia a todo lo que es susceptible de definirse.” (*Idem* §10, p. 32). Como ejemplo de esta clase de términos pueden citarse: «caballo» y «quimera». Es importante notar que esta caracterización de los términos complejos revela una estrecha relación, en el pensamiento de Moore, entre las nociones de descripción y definición. Según Moore, cuando se da una *descripción* del objeto al que se refiere un término complejo como «caballo» (entendiendo esta descripción, como se dijo, como una relación de las partes que componen a tal objeto), en realidad se está dando una *definición* de dicho término. De ahí que se diga que: “la principal acepción de «definición» es aquella por la que una definición declara cuáles son las partes que invariablemente componen un cierto todo.” (*Idem* §10, p. 32). De este modo, si, como afirman los

Naturalistas, es posible definir «bueno», entonces esta palabra es de la misma clase de «caballo» y «quimera», es decir, es un término complejo.

Finalmente, una cuarta afirmación del Naturalismo es que es posible probar algunos Juicios Morales (en adelante, simplemente JM). Por JM entiendo un enunciado en el que se dice de una determinada acción o persona que posee cierta propiedad moral. En ese sentido, enunciados éticos como las definiciones que proponen los Naturalistas no son JM, aún cuando en ellos se empleen términos morales; ya que tales enunciados no están referidos a acciones o personas concretas, sino que buscan develar la naturaleza de la bondad misma. Ahora, volviendo a la cuarta afirmación del Naturalismo, ella es una consecuencia de la primera y de la segunda afirmación, esto es, de la idea de que es posible definir «bueno» y de la idea de que existen enunciados éticos analíticos. En efecto, dado que en toda definición el *definiendum* y el *definiens* son expresiones sinónimas (por consiguiente, lógicamente equivalentes e intercambiables entre sí); la existencia de una definición de «bueno» hace posible una *reducción* del vocabulario moral al vocabulario en términos del cual se ha definido dicha palabra. En consecuencia, la investigación ética, entendida como una investigación orientada a determinar la corrección o incorrección de ciertos JM, debe realizarse al interior de la disciplina cuyo vocabulario específico contenga los términos que han sido empleados para definir «bueno». Esta investigación debe llevarse a cabo de acuerdo con los métodos propios de la disciplina en cuestión y debe apuntar a determinar si las acciones o personas sobre las que versa un determinado JM poseen o no las propiedades a las que se hace referencia mediante los términos que han sido usados para definir «bueno». Así, desde esta perspectiva, la prueba de cualquier JM debe tener como primera premisa la definición de «bueno» y debe desenvolverse del siguiente modo:

«Bueno» es «X»  
a es X  
Por lo tanto, a es buena,

Donde «X» es un término propio de la disciplina a la cual se ha reducido el vocabulario moral y a es una acción o persona cualquiera. Como hemos dicho, la investigación ética debe buscar determinar el valor de verdad de la segunda premisa, «a es X», para lo cual debe valerse de los métodos propios de la disciplina cuyo vocabulario específico contenga el término «X».

Tenemos, así, que la doctrina conocida como Naturalismo debe comprometerse con la verdad de cuatro afirmaciones fundamentales, a saber:

- (1) Es posible definir el término «bueno».
- (2) Hay proposiciones éticas analíticas.
- (3) «Bueno» es un término complejo.
- (4) Es posible probar algunos JM.

Sin embargo, es importante notar que esta caracterización aún no permite apreciar lo específico del Naturalismo. Como Moore ha señalado, teorías éticas como la de los estoicos, la de Spinoza, la de Kant y la de Hegel, las cuales evidentemente no son Naturalistas, de todos modos se comprometen con la verdad de las afirmaciones (1)-(4). Tales teorías se caracterizan por definir la palabra «bueno» empleando términos como «Substancia Absoluta» (Spinoza), «Reino de los Fines» (Kant) o «Verdadero Yo» (Hegel); esto es, empleando términos que hacen referencia a una *realidad suprasensible*, distinta del mundo natural. Por tal razón, Moore denomina “Metafísicas” a las teorías de este tipo y las distingue de las Naturalistas (Cf. *Idem* §67, p. 140). Pero, entonces, ¿qué es lo

específico del Naturalismo? ¿Qué permite distinguir a las teorías Naturalistas de estas teorías Metafísicas? Para aproximarnos al Naturalismo, aún debemos dar un paso adicional, aún debemos responder a estos interrogantes.

Según Moore, la diferencia entre el Naturalismo y una perspectiva Metafísica radica en que, el primero, pero no el segundo, debe comprometerse con la verdad de una tesis adicional, a saber, que «bueno» debe definirse empleando *términos naturales*. Pero, ¿qué es un término natural? Consideremos algunas afirmaciones de Moore, acerca de la noción de “natural”:

[...] esta cosa, con referencia a la cual se define bueno, podría ser [...] lo que puedo llamar un objeto natural –algo cuya existencia se está de acuerdo en que es un objeto de la experiencia– [...]. (*Idem* §25, p. 62)

Con «naturaleza», he querido y quiero referirme al contenido de las ciencias naturales y también de la Psicología. Puede decirse que incluye todo lo que ha existido, existe y existirá en el tiempo. (*Idem* §26, p. 64)

Estos pasajes revelan que, para Moore, un objeto natural es aquel que puede conocerse a través de los sentidos, esto es, del cual puede tenerse algún tipo de experiencia, y cuya explicación es algo propio del dominio de las ciencias empíricas. Es importante resaltar que, de acuerdo con esta caracterización, que un objeto sea natural no implica necesariamente que sea un objeto físico. Según Moore, cualquier objeto propio del dominio de ciencias como la Psicología y la Sociología es natural. Esto se debe a que estas ciencias son empíricas; además, estudian fenómenos que, aun cuando no son físicos, existen en el tiempo y hacen parte de la estructura del mundo (en otras palabras, las percepciones, las creencias, los deseos, las revoluciones políticas y sociales, etc. existen en un tiempo determinado y hacen parte de nuestra ontología). Podemos decir, entonces, que un término es natural cuando se refiere a un objeto natural, en el sentido que acabamos de especificar. El Naturalismo, en consecuencia, debe comprometerse con la tesis, bastante polémica, de que la bondad es un objeto natural, el cual debe ser investigado en el marco de las ciencias empíricas. Esto se puede apreciar claramente, al considerar los ejemplos que hemos mencionado de Naturalismo: las Teorías del Interés y al Utilitarismo. En efecto, en la teoría de Hobbes y en el Utilitarismo, «bueno» se define en términos que son propios del vocabulario de la Psicología, a saber, «deseo» y «placer», respectivamente; mientras que en la teoría de Hume, «bueno» se define en términos que pueden hacer parte del vocabulario de la Sociología, a saber, «aprobado por la mayor parte de la gente».

Adicionalmente, el compromiso con la verdad de esta afirmación adicional tiene como consecuencia una concepción específica de la tesis (4), es decir, de cómo debe llevarse a cabo la prueba de un JM. Habíamos dicho que, desde la perspectiva que estamos considerando, al probar un JM específico, la investigación debe concentrarse en determinar el valor de verdad de la segunda premisa del argumento, esto es, «a es X»; lo cual debe hacerse de conformidad con el método de investigación propio de la disciplina cuyo vocabulario específico contiene a «X». Como acabamos de señalar, desde el punto de vista del Naturalismo, tal disciplina siempre será una ciencia empírica. En consecuencia, la determinación del valor de verdad de la premisa en cuestión siempre ha de hacerse valiéndose de los métodos propios de este tipo de ciencias. Con lo cual, según el Naturalismo, la investigación ética tendrá un carácter esencialmente empírico.

Ahora si podemos dar una caracterización completa del Naturalismo. Como hemos visto, esta doctrina está constituida por las afirmaciones (1)-(4), las cuales con comunes a

las teorías Metafísicas, y por las siguientes afirmaciones adicionales, las cuales le dan su carácter específico:

- (5) «Bueno» debe definirse empleando términos naturales.
- (6) Toda prueba de un JM debe valerse de los métodos de las ciencias empíricas.

Con esta caracterización a la mano, ahora debemos preguntarnos: ¿cuál es el error conceptual en que incurren los Naturalistas? ¿en qué consiste eso que Moore denomina la Falacia Naturalista? Trataremos de responder a estas inquietudes a continuación.

Para Moore, la Falacia Naturalista es un caso particular de una falacia más general, a la que, siguiendo a Frankena, podemos denominar “Falacia de Definición” (1939, p. 471). En esta falacia incurre todo aquel que intenta definir un término simple, esto es, todo aquel que intenta definir un término cuya naturaleza es tal que, como hemos visto, realmente es indefinible. Quien pretende haber definido un término simple tiende a tratar a éste como sinónimo de la expresión que empleó para definirlo; de modo que, tiende a pensar que ambos, el supuesto *definiendum* y el supuesto *definiens*, son términos o expresiones que se refieren a un mismo objeto. Por ello, esta falacia también puede caracterizarse como el error en que incurre quien, confundiendo dos objetos diferentes, los trata como si el primero fuera idéntico al segundo (Cf. Moore 1903, §12, p. 36). Ahora, según Moore, «bueno» es un término simple, un término de la misma clase de «amarillo», el cual se refiere a un objeto simple de la experiencia. Debido a que los Naturalistas no se dan cuenta de este hecho, ellos se comprometen con la idea de que las afirmaciones (1)-(3) son verdaderas, cuando en realidad ellas son falsas, y por ello comenten la Falacia de Definición. Adicionalmente, es importante anotar que, quien cae en esta falacia, inmediatamente se ve obligado a admitir la afirmación (4), a saber, que es posible probar algunos JM; ya que, como hemos visto, la noción de una prueba de un JM se basa en la idea de que existe una definición de «bueno». Según esto, la Falacia de Definición es un error que está a la base de toda teoría en la que las afirmaciones (1)-(4) sean verdaderas; por consiguiente, es un error que está a la base tanto de las teorías Naturalistas como de las teorías Metafísicas.

Ahora, para Moore, «bueno» no sólo es un término simple; también es un término que tiene un carácter especial, diferente al que poseen los términos naturales. Esto quiere decir que, según él, la palabra «bueno» se refiere a un objeto simple de la experiencia que posee un carácter *especial*, esto es, que no pertenece a ninguna de las categorías que empleamos para clasificar a los objetos naturales. De este modo, la Falacia Naturalista puede concebirse como el error en que incurre quien, creyendo que es posible definir la palabra «bueno», trata de definirla empleando términos naturales. Para ponerlo de otro modo, quien no sólo admite la verdad de las afirmaciones (1)-(4), sino que también admiten la verdad de las afirmaciones (5)-(6), comete la Falacia Naturalista. Así, lo específico de ella, aquello que permite distinguirla de la simple Falacia de Definición, es que quien la comete se compromete con la idea de que el objeto al que nos referimos con la palabra «bueno», esto es, la bondad misma, puede identificarse con un objeto natural:

Quando alguien confunde dos objetos naturales entre sí definiendo uno por otro, si en un suponer él, que es un objeto natural, se confundiera a sí mismo con «contento» o con «placer», que son otros; entonces no habría motivo para designar esta falacia como naturalista. Pero si confunde «bueno», que no es un objeto natural en el mismo sentido, con cualquier otro objeto natural, entonces sí que habría que denominarlo una falacia naturalista [...]. (Moore 1903, §12, p. 36)

Este pasaje muestra claramente por qué Moore considera que la Falacia Naturalista es un caso especial no es más que un caso especial de la Falacia de Definición.

Lo que hemos dicho hasta acá, acerca del Naturalismo y del error que Moore ve en él, nos permite comenzar a hacernos una idea de en qué consiste el Intuicionismo. Como dijimos al inicio de este capítulo, Moore concibe al Intuicionismo como una doctrina capaz de superar los problemas del Naturalismo, es decir, como una doctrina que no comete la Falacia Naturalista. Según esto, un primer aspecto de tal doctrina es que, en ella, las afirmaciones (1)-(4) se consideran falsas; mientras que las siguientes se consideran verdaderas:

- (7) No es posible definir el término «bueno».
- (8) Toda proposición ética es sintética.
- (9) «Bueno» es un término simple.
- (10) Es posible probar algunos JM

Podemos considerar a (9) como una de las afirmaciones fundamentales del Intuicionismo, de la cual se derivan las demás. En efecto, hemos visto que, para Moore, sólo los términos complejos son definibles; de modo que, si (9) es verdadera, entonces (7) también lo será. Adicionalmente, si no existe una definición de «bueno», esto es, si (7) es verdadera; entonces no existe un enunciado ético en el que el sujeto y el predicado sean expresiones sinónimas y, por consiguiente, no existen enunciados éticos analíticos. Así, si (7) es verdadera, (8) también lo será. De acuerdo con (8), enunciados como el Principio Utilitarista, los cuales son considerados como definiciones por los Naturalistas, en realidad expresan una relación entre dos objetos, en este caso, entre la bondad y el placer. Por tal razón, según Moore, estos enunciados deben ser considerados como *Principios Éticos Sustantivos* y deben desempeñar un rol central en lo que él denomina el “ideal” de la ciencia ética, esto es, en el desarrollo de la ética práctica. Por su parte, la afirmación (10) es una de las tesis más importantes en la concepción mooreana del Intuicionismo. Como vimos, la idea de que es posible probar algunos JMs se deriva de la creencia de que es posible definir la palabra «bueno», ya que toda prueba de un JM debe tomar como primera premisa a dicha definición. Así, si (7) es verdadera, entonces (10) también lo será. Ahora, lo interesante con esta afirmación es que detrás de ella hay una concepción bastante particular acerca de qué es lo que permite determinar la corrección o incorrección de un JM. En efecto, Moore afirma:

[...] desearía que se observara que, cuando llamo *intuiciones* a tales proposiciones [i.e., a los JM], simplemente quiero apuntar a que no pueden probarse (*Idem* Prólogo, p. 20)

Creo que, al llamar “intuiciones” a los JM, Moore quiere resaltar algo como lo siguiente. En vista de que él piensa que no es posible probar ningún JM y que la bondad es un objeto simple de nuestra experiencia; desde su punto de vista, siempre que queramos saber si un JM determinado es correcto o incorrecto, a lo único a lo que podemos apelar es a nuestra *experiencia moral* particular. Así, puede decirse que, si bien el Intuicionismo se caracteriza por negar la tesis de que es posible definir el término «bueno»; lo que realmente lo distingue de las posiciones Naturalistas y Metafísicas es que él involucra una concepción experiencial de nuestras prácticas morales. Más adelante veremos que una de las cosas que, justamente, ha motivado muchas de las críticas que se le han hecho al Intuicionismo de Moore, es que en sus planteamientos no hay una concepción muy clara de en qué consiste esa experiencia moral que él parece colocar a la base de nuestros JMs.

El segundo aspecto del Intuicionismo tiene que ver propiamente con el carácter específico de la Falacia Naturalista, esto es, con aquello que la distingue de la Falacia de Definición. Según esto, el Intuicionismo se caracteriza por rechazar las afirmaciones (5)-(6), las cuales constituyen de manera específica al Naturalismo. En contraste, esta doctrina está constituida por la siguiente afirmación:

(11) «Bueno» no es un término natural.

Es importante notar que esta afirmación no se sigue directamente de la negación de (5)-(6). Como vimos, el Intuicionismo comienza por negar las afirmaciones (1)-(4), afirmando que toda doctrina que las admite incurre en la Falacia de Definición. De esto se sigue inmediatamente la negación de (5)-(6); puesto que, si no es posible definir «bueno», entonces es trivialmente verdadero que no es posible definir esta palabra empleando términos naturales. Sin embargo, esto no implica que toda forma de Naturalismo es errónea. Hasta ahora, hemos empleado la denominación Naturalismo para referirnos a una doctrina que, propiamente hablando, es una forma de *Naturalismo Definicional*. Pero no todas las formas de Naturalismo son Definicionales. El Naturalista perfectamente puede aceptar las consideraciones de Moore acerca de la Falacia de Definición, y no obstante, continuar afirmando que «bueno» es un término natural. En tal caso, el Naturalismo tendría que comprometerse con una tesis como la siguiente: «bueno», al igual que «amarillo», es un término *simple y natural*, esto es, un término que se refiere a un objeto simple de la experiencia, el cual pertenece al dominio de alguna de las ciencias naturales. Es evidente que el propio Moore llegó a contemplar la posibilidad de que el Naturalismo adopte una forma no Definicional; esto puede apreciarse en el siguiente pasaje:

Con respecto a las razones por las que bueno no debe considerarse un objeto natural, pueden discutirse en otro lugar. En este momento, es suficiente prestar atención a esto, que incluso si fuera un objeto natural, ello no alteraría la naturaleza de la falacia ni disminuiría en lo más mínimo su importancia. (*Idem*, p. 36)

Aquí Moore se refiere a lo que propiamente hemos llamado la Falacia de Definición. De tal modo, en este pasaje Moore resalta que, incluso si admitimos que «bueno» se refiere a un objeto natural, dado que tal objeto es simple, de todos modos quien intenta definir la palabra «bueno» incurre en un error. Por tal razón, es erróneo considerar que la afirmación (11) se deriva trivialmente de (7)-(10). Por consiguiente, (11) es una tesis adicional del Intuicionismo, una tesis que es independiente de las consideraciones acerca de la simpleza y el carácter indefinible de la palabra «bueno».

Para terminar, es importante resaltar que, desafortunadamente, a pesar de que Moore contempla la posibilidad de que el Naturalismo sea no Definicional, él tiende a tratar a todas las teorías Naturalistas como formas de Naturalismo Definicional. Por esta razón, él piensa, como hemos visto, que la Falacia Naturalista simplemente es un caso particular de la Falacia de Definición. En realidad, estas dos falacias corresponden a dos tipos de error diferentes, conceptualmente independientes uno del otro. Esto se puede apreciar, si nos percatamos de que es posible cometer la Falacia Naturalista sin caer en la Falacia de Definición, tal como ocurre cuando se sostiene una forma de Naturalismo no Definicional; y a la inversa, es posible cometer la Falacia de Definición sin caer en la Falacia de Naturalista, tal como ocurre cuando se sostiene una teoría Metafísica. De igual modo, en la medida en que Moore no distingue claramente estas dos falacias, tiende a pensar que las dos tesis fundamentales del Intuicionismo, a saber, las afirmaciones (9) y (11), están estrechamente relacionadas y pueden justificarse mediante un único argumento. Como



veremos, a continuación, esta idea es la que ha permitido plantear las objeciones más serias al Intuicionismo de Moore.

## 1.2. El Argumento de la Pregunta Abierta y el Internalismo

Las consideraciones del numeral anterior nos dejan ver que, a los ojos de Moore, la Falacia Naturalista es la mejor herramienta para resolver la polémica entre Naturalismo e Intuicionismo. Según él, el Naturalismo es inadmisibile, ya que incurre en la Falacia Naturalista; mientras que el Intuicionismo es la mejor teoría acerca del lenguaje moral, porque él no comete dicha falacia. Así, a los ojos de Moore, la Falacia Naturalista debe ser empleada para formular el criterio definitivo, en virtud del cual se ha de determinar si una teoría metaética es correcta o no. Sin embargo, W.K. Frankena ha cuestionado duramente esta estrategia. Como vimos, Moore no distingue claramente entre la Falacia Naturalista y la Falacia de Definición. Por tal razón, sin darse cuenta, al criticar al Naturalismo, él se vale propiamente de la Falacia de Definición, no de la Falacia Naturalista. En efecto, como se dijo, el principal problema que Moore encuentra en las teorías Naturalistas es que ellas no logran apreciar el carácter *simple e indefinible* del objeto al que nos referimos con la palabra «bueno». Según Frankena, podemos conceder que la Falacia de Definición es un error conceptual (algo que, de todos modos, debe quedar abierto a discusión); y a partir de ello, podemos decir que el Naturalismo Definicional es insostenible, por cometer precisamente ese error. Sin embargo, de esto no se sigue que el Intuicionismo es verdadero; puesto que una tesis fundamental de él es que «bueno» se refiere a un objeto *especial*, de una clase diferente a la de los objetos naturales, y hasta ahora no se ha dicho nada que nos lleve a pensar que esta tesis es verdadera. Debido a que Moore confunde a la Falacia de Definición con la Falacia Naturalista; él considera que, al atribuirle la primera al Naturalista, inmediatamente puede pasar a atribuirle la segunda. Sin embargo, es evidente que para dar este paso adicional, esto es, para poder afirmar que el Naturalista comete un error al sostener que «bueno» se refiere a un objeto natural, es necesario probar de ante mano que, efectivamente, esta palabra se refiere a un objeto que pertenece a una clase diferente a la de los objetos naturales. En otras palabras, para atribuirle la Falacia Naturalista a cualquier teoría, antes es necesario mostrar que el Intuicionismo es verdadero. A la luz de estas consideraciones, Frankena acusa a Moore de cometer una *petición de principio*: para darle sustento al Intuicionismo, él afirma que el Naturalismo es erróneo por incurrir en la Falacia Naturalista; pero la atribución de la Falacia Naturalista sólo es posible, si el Intuicionismo es verdadero. En consecuencia, a diferencia de la Falacia de Definición, esta falacia no puede ser usada como una herramienta para resolver la polémica entre Naturalismo e Intuicionismo. Por el contrario, sólo cuando dicha polémica haya sido resuelta a favor del Intuicionismo, podrá acusarse al Naturalista de cometerla. El Intuicionismo, entonces, debe justificarse a partir de consideraciones adicionales, que no involucren referencia alguna a la Falacia Naturalista (Cf. 1939, pp. 471-4).

Quien conoce la obra de Moore puede pensar que esta crítica no necesariamente lleva a rechazar los planteamientos de éste acerca del Intuicionismo. Esto debido a que, de todos modos, Moore sí parece disponer de un argumento adicional, que le permite sostener la idea de que «bueno» se refiere a un objeto de carácter especial. Tal argumento es conocido como el *Argumento de la Pregunta Abierta* (en adelante, APA), el cual, siguiendo a Alexander Miller, puede esquematizarse del siguiente modo. En primer lugar, debemos partir de la tesis del Naturalista:

- (a) El predicado «bueno» es sinónimo de, o analíticamente equivalente a, un predicado natural «N».

De esta afirmación se sigue:

- (b) Parte del significado de la afirmación «x es N» está constituido por la afirmación «x es bueno».

Ahora, si (a) y (b) son verdaderas, entonces:

- (c) Alguien que pregunta de un x que es N si también es bueno, exhibe una confusión conceptual.

Sin embargo:

- (d) Dada una propiedad natural N, siempre es una *pregunta abierta* si un x que es N también es bueno. Esto quiere decir que siempre es significativo (no un sinsentido) preguntar, de un x que es N, si también es bueno: preguntar «¿un x que es N, también es bueno?» No exhibe ninguna confusión conceptual.

Así, de (c) y (d), por reducción al absurdo, se deduce:

- (e) No es el caso que el predicado «bueno» es sinónimo de, o analíticamente equivalente a, un predicado natural «N».

Finalmente:

- (f) La propiedad de «ser bueno» no puede ser idéntica por necesidad conceptual (i.e., por definición) a la propiedad de «ser N» (Cf. Miller 2003, pp. 13-4).

En términos generales, la idea de Moore es que si el predicado «bueno» es por definición equivalente a un predicado natural «N», entonces, del mismo modo en que no tiene sentido preguntar ¿un x que es bueno, también es bueno?, tampoco debería tener sentido preguntar ¿un x que es N, también es bueno? La primera pregunta es una pregunta cerrada, esto es, una pregunta carente de sentido. Sin embargo, en vista de que esta última pregunta es una pregunta abierta, es decir, una pregunta que sí tiene sentido formular, no es posible sostener el Naturalismo (Cf. Moore 1903, §13, pp. 38-9).

Creo que hay dos razones por las cuales es erróneo pensar que APA aporta las consideraciones adicionales que, según Frankena, son necesarias para darle sustento al Intuicionismo. La primera de ellas es que, tal como está puesto, el argumento apunta a criticar al Naturalismo Definicional, esto es, apunta a resaltar el hecho de que todo aquel que pretenda dar una definición de «bueno» en términos naturales necesariamente deberá enfrentar el problema de la pregunta abierta. Algo que muestra claramente que lo que está en cuestión aquí es la indefinibilidad del término «bueno», es el hecho de que este argumento puede modificarse, de un modo tal que sirva para cuestionar a las teorías Metafísicas. En efecto, sólo con que en las premisas reemplacemos a la variable «N», la cual alude a cualquier término natural, por una variable «M», que aluda a cualquier término metafísico; podremos formular el argumento, de un modo tal que pueda aplicarse a cualquier teoría Metafísica (Miller 2003, pp. 14-5). Así, el argumento simplemente sirve para probar la tesis de que «bueno» es un término simple e indefinible; en ese sentido, el argumento sólo permite sustentar la idea de que el Naturalismo Definicional comete la Falacia de Definición. Sin embargo, este resultado aún es compatible con la idea de que «bueno» es un término simple y natural, el cual no puede definirse empleando otros

términos naturales. Así, APA no parece afectar en nada al Naturalismo no Definicional. De todas formas, un defensor de Moore podría alegar que el argumento muestra que «bueno» tiene *rasgos semánticos especiales*, en virtud de los cuales, él puede distinguirse de cualquier término natural. En ese sentido, APA no sólo probaría que «bueno» es un término simple e indefinible; sino que también mostraría que es un término especial, puesto que sus rasgos semánticos impiden identificarlo con un término natural cualquiera. La segunda razón, por la que considero que APA no aporta las consideraciones adicionales necesarias para fundamental al Intuicionismo, tiene que ver con esta réplica. De acuerdo con ella, lo que prueba que «bueno» tiene rasgos semánticos especiales, es el hecho de que la pregunta que se formula en el paso (d) del argumento siempre es una pregunta abierta. Que esta pregunta siempre sea abierta, sólo puede explicarse por el hecho de que ningún predicado natural posee las propiedades semánticas propias de un predicado moral como «bueno». Sin embargo, como Miller ha señalado, si APA es visto de este modo, entonces Moore nuevamente está cometiendo una petición de principio, análoga a la que Frankena había denunciado. En efecto, ¿por qué tendríamos que decir que, para el Naturalista que considera que «bueno» es un término natural, la pregunta del paso (d) siempre es una pregunta abierta? Parece que, para que podamos decir que, en efecto, ella siempre es una pregunta abierta, de antemano debemos admitir la idea de que «bueno» es un término que posee rasgos semánticos especiales. Pero esta última tesis es, justamente, la que se pretende probar con el argumento. Así, al emplear APA para tratar de sustentar al Intuicionismo, Moore cae nuevamente en una petición de principio (Miller 2003, pp. 15-6).

Lo anterior no implica que no haya ningún recurso adicional, del que pueda valerse el Intuicionista para tratar de sustentar la idea de que «bueno» es un término especial. De hecho, en el contexto la discusión acerca de cuál es la diferencia entre el Intuicionismo y el Naturalismo Definicional, el propio Frankena algunas observaciones bastante sugerentes:

What underlies this difference of opinion is that the intuitionists claim to have at least a *dim awareness of a simple unique quality* or relation of goodness or rightness which appears in the region which our ethical terms roughly indicate, whereas the definists claim to have *no awareness of any such quality or relation* in that region, which is different from all other qualities and relations which belong to the same context but are designated by words other than 'good' and 'right' and their obvious synonyms (1939, p. 474. Énfasis mío)

Y más adelante continúa:

The issue, then, is one of *inspection or intuition*, and concerns the *awareness or discernment of qualities and relations*. That is why it cannot be decided by the use of the notion of a fallacy. (*Idem*, p. 474. Énfasis mío)

Creo que estos pasajes apuntan a algo crucial, a saber, que, a fin de sustentar su idea de que el término «bueno» posee rasgos semánticos especiales, el Intuicionista debe recurrir, no a la noción de una Falacia Naturalista, sino a un *examen de la experiencia* que tenemos cuando hacemos JMs. Según esto, el Intuicionismo debe fundamentarse principalmente en consideraciones fenomenológicas. A partir de esto, Stephen Darwall, Allan Gibbard y Peter Railton han sugerido que, para no caer en una petición de principio al formular APA, el Intuicionista debe resaltar el hecho de que el lenguaje moral tiene un carácter *esencialmente práctico*, o en otras palabras, el lenguaje moral tiene como función principal guiar nuestras acciones. Esto es algo de lo cual nos podemos dar cuenta, examinando la experiencia que tenemos en la vida cotidiana al momento de hacer JMs. En efecto, en condiciones normales, siempre que un individuo A juzga sinceramente que

«x es buena», tendrá la motivación para hacer x, cuando las circunstancias se lo permitan. Habiendo juzgado que «x es buena», si las circunstancias son propicias y resulta que A no hace x; entonces, lo más probable es que nos sorprendamos ante el comportamiento de A y tratemos de explicarlo diciendo, o bien que A tiene razones mucho más fuertes para no hacer x, o bien que en realidad A no juzgó sinceramente que «x es buena», o bien que en realidad A no comprende el significado de la palabra «bueno». Este hecho indica que hay un *vínculo interno* entre nuestro lenguaje moral y nuestras motivaciones para actuar. En otras palabras, esto es un indicio de que hay un elemento en el contenido de nuestros términos morales, tal que, cuando los empleamos en nuestros juicios, éstos últimos se constituyen en motivos o razones para actuar. A esta posición se le conoce como Internalismo. En el capítulo 4 nos ocuparemos en detalle de esta doctrina y la defenderemos de las objeciones que algunos opositores han presentado contra ella. Hasta que lleguemos allí, daremos por sentado que el Internalismo es verdadero, basándonos en las observaciones fenomenológicas que acabamos de hacer. Por ahora, lo que debemos tener en mente es que, gracias al Internalismo, podemos decir que el rasgo especial de términos morales como «bueno» es que ellos poseen un vínculo interno con nuestras motivaciones para actuar, vínculo que los términos naturales no parecen poseer. Y si esto es así, entonces, el Intuicionista tiene a la mano las consideraciones adicionales, independientes de la Falacia Naturalista y del APA, que le permiten sustentar su tesis de que «bueno» es un término especial, cuyos rasgos semánticos difiere de los rasgos de los términos naturales (Darwall, Gibbard y Railton 1992, pp. 118-9)

Una observación final. Hasta acá, hemos señalado algunos inconvenientes que hay en los planteamientos de Moore acerca del Intuicionismo y hemos sugerido cómo estos inconvenientes pueden superarse, a la luz de ciertas afirmaciones de autores como Frankena, Miller, Darwall, Gibbard y Railton. Sin embargo, en los *Principia Ethica* no hay muchos elementos que sugieran que Moore estaría dispuesto a abrazar al Internalismo, como la doctrina capaz de explicar en qué consiste el carácter especial del lenguaje moral. Por tal razón, no estoy seguro que las afirmaciones que haremos a continuación, acerca del Intuicionismo, puedan aplicarse a éste, tal como Moore lo concibió. Con esto en mente, en adelante emplearé la denominación “Intuicionismo”, no para referirme específicamente a doctrina desarrollada por Moore; sino para referirme a la doctrina general, que requiere del Internalismo para explicar en por qué «bueno» no es un término natural.

### 1.3. Epistemología y Metafísica Especiales

Hemos visto que el Intuicionista puede escapar a muchas de las críticas que se le han hecho a los planteamientos de Moore, gracias a una descripción de nuestra experiencia moral, que muestre cómo el Internalismo permite dar cuenta del carácter especial del lenguaje moral. Sin embargo, esta estrategia parece generar muchos más problemas, que los que resuelve. El propósito de este numeral es considerar estos problemas.

Un aspecto que aún no hemos resaltado del Intuicionismo es que, al igual que el Naturalismo, él es una teoría Cognitivist. Esto quiere decir que, para los Intuicionistas, los JMs tienen contenido representacional; por consiguiente, ellos pueden ser evaluados en términos de verdad y falsedad. Según esto, un JM que es verdadero, no sólo le da un motivo para actuar a quien lo emite; sino que, además, le da algún tipo de *conocimiento*. En varios pasajes de los *Principia*, podemos ver que, efectivamente, Moore se compromete con una visión Cognitivist del lenguaje moral:

En la mayoría de los casos, cada vez que empleamos términos como «virtud», «vicio», «obligación», «correcto», «debe», «bueno», «malo», estamos haciendo con toda seguridad

juicios éticos; y, si queremos discutir su *verdad*, tendremos que debatir un aspecto de la Ética. (Moore 1903, §1, p. 23. c)

Debo insistir, pues, en que no sólo es competencia de la Ética la obtención de unos resultados *verdaderos*, sino también la búsqueda de unas razones válidas para tales resultados. *El objeto directo de la Ética es el conocimiento y no la práctica*. De modo que, por correctos que puedan ser los principios prácticos a los que se llegue, cualquiera que se sirva de la falacia naturalista no cumple sin duda este primer objeto. (*Idem*, §14, p. 43. Énfasis mío)

Es de esperar, entonces, que cualquier teoría Intuicionista, inspirada en los planteamientos de Moore, admita estas afirmaciones. Ahora, quienes sostienen el Cognitivismo, consideran que existe una instancia, independiente de los intereses particulares de los hablantes, en virtud de la cual es posible evaluar la corrección o incorrección de un JM. Gracias a esto, es posible criticar las opiniones morales de otras personas, cuando consideramos que los Juicios mediante los cuales ellas se expresan son erróneos. Así, el Cognitivismo parece una posición necesaria, para dar cuenta de la idea de que el lenguaje moral es objetivo. En el capítulo 3, discutiremos en detalle lo relativo a esta idea. Por lo pronto, simplemente me interesa resaltar que el Intuicionismo es una teoría de corte Cognitivista, que pretende explicar la objetividad del lenguaje moral, atribuyéndole condiciones de verdad a los JMs.

Sin embargo, algunos autores han señalado que, la pretensión del Intuicionismo de explicar la objetividad del lenguaje moral queda completamente frustrada, debido al rol que, en esta teoría, desempeña la experiencia moral. El problema en relación a la experiencia moral surge, debido a que el Intuicionista nunca especifica con claridad en qué consiste ella. Para intentar darle sentido a los planteamientos Intuicionistas, autores como A.J. Ayer y R.M. Hare han propuesto entender a la experiencia moral como ese *sentimiento* que con frecuencia acompaña a nuestras apreciaciones morales; como, por ejemplo, el sentimiento de indignación tenemos, cuando juzgamos que cierta acción es injusta, o el sentimiento de culpa tenemos, cuando juzgamos que hemos actuado mal. El problema es que este tipo de sentimientos nunca están ligados de manera necesaria a la representación de nuestras acciones. Por el contrario, el vínculo entre un sentimiento específico y la representación de una acción varía notablemente, dependiendo de la educación de una persona y del contexto histórico y social al que pertenece. Así, por ejemplo, en la cultura occidental solemos tener un sentimiento de desagrado e indignación al representarnos acciones como la ablación del clítoris que se le practica a algunas mujeres en países africanos y asiáticos; sin embargo, es muy probable que en estos países las personas no experimenten este sentimiento, o que experimenten un sentimiento más positivo, al representarse este tipo de prácticas. De este modo, si decimos que un JM como «x es buena» es verdadero, cuando la representación de x está acompañada del sentimiento que asociamos a la palabra «bueno»; entonces, este Juicio será verdadero para algunas personas y falso para otras personas. En consecuencia, de acuerdo con esta visión de la experiencia moral, el Intuicionismo es incapaz de dar cuenta de la objetividad del lenguaje moral; ya que no dispone de una instancia independiente del pensamiento y de los intereses particulares de los hablantes, que le permita determinar cuando un JM es verdadero y cuando es falso (Cf. Ayer 1952, pp. 123-4; Hare 1997, pp. 93-9).

Algunos autores, como J.L. Mackie, han sugerido que, para salvar al Intuicionismo de este tipo de críticas, es necesario comprometerlo con la tesis de que poseemos una *facultad moral especial*. Así, la experiencia moral puede ser vista como el producto del ejercicio de esa facultad. Pero adicionalmente, para que sus pretensiones de objetividad no se vean frustradas, el Intuicionismo debe comprometerse, adicionalmente, con una tesis metafísica, a saber, que las propiedades morales hacen parte de la estructura de la

realidad, que son reales e independientes de cualquier punto de vista subjetivo. Ahora, dado que para el Intuicionismo los términos morales tienen un carácter especial, el cual se explica a través del Internalismo, esas propiedades igualmente deben tener un carácter especial. Ellas deben ser concebidas como *propiedades prescriptivas*, capaces de guiar nuestras acciones; y por ello, pertenecientes a una clase diferente a la que pertenecen las propiedades naturales. Según Mackie, tales propiedades deberán ser un análogo a la Idea Platónica del Bien; ya que el simple conocimiento de ellas debe ser suficiente para motivar las acciones del agente. Así, parece que el Intuicionismo también debe suscribir una *metafísica especial*; de modo que, nuestros JMs puedan ser vistos como el ejercicio de nuestra facultad moral especial, la cual nos permite conocer una realidad moral especial. La semántica del lenguaje moral, de acuerdo con esta posición, debe explicarse mediante esta epistemología y esta metafísica especiales. «Bueno», según esto, debe concebirse como un término que usamos para referirnos a una propiedad prescriptiva real, a saber, la bondad; la cual podemos conocer gracias a nuestra facultad moral especial (Cf. Mackie 1977, pp. 42-6).

No hay que hacer un examen muy cuidadoso de esta propuesta, para darse cuenta de su implausibilidad. De hecho, si esta es la única alternativa que tiene el Intuicionismo para dar cuenta, simultáneamente, de sus pretensiones de objetividad y del Internalismo; entonces, claramente, no es una buena alternativa para tratar de explicar la semántica del lenguaje moral. Esta es la razón por la que muchos autores contemporáneos consideran al Intuicionismo como una posición completamente desacreditada. Ahora, no estoy seguro de que esta sea una buena interpretación del Intuicionismo; en particular, porque creo que es posible darle sentido a la noción de experiencia moral sin recurrir a sentimientos subjetivos ni a facultades especiales. En el capítulo 3 esbozaré una teoría que se aproxima bastante al Intuicionismo y que, respaldada por numerosas investigaciones de carácter empírico, puede especificar una noción diferente de experiencia moral. De momento, es importante resaltar que, a partir de estas consideraciones en torno al Intuicionismo, ahora podemos formular con precisión el problema que deben enfrentar, quienes pretenden desarrollar una teoría acerca de la semántica del lenguaje moral. Como acabamos de ver, lo que lleva al Intuicionismo a sostener una epistemología y una metafísica especiales, es su compromiso con dos tesis fundamentales acerca de la moralidad, a saber:

**Objetividad:** nuestros JMs son objetivos, es decir, son evaluables como correctos o incorrectos, bajo criterios independientes de cualquier punto de vista subjetivo. Cuando son correctos, ellos son válidos para todo hablante, independientemente de que él los acepte o no.

**Internalismo:** existe una conexión interna entre nuestros JMs y nuestras motivaciones para actuar.

La tesis de la Objetividad parece llevar necesariamente al Cognitivismo. De modo que, el problema, en últimas, es cómo sostener simultáneamente al Cognitivismo y al Internalismo, sin caer en una teoría como el Intuicionismo que acabamos de considerar. Michael Smith (1994) ha denominado a este problema el *Problema Moral* y ha sostenido, creo que correctamente, que él constituye el problema central de la metaética. En lo que sigue examinaré un poco más este problema, a fin de revelar un supuesto que hay tras su formulación y que es determinante para comprender exactamente a qué nos estamos enfrentando.

## 1.4. El Problema Moral

Según Smith, el Problema Moral surge del hecho de que hay tres tesis relevantes para la comprensión del lenguaje moral, a las cuales aparentemente no podemos renunciar; pero que, no obstante, parecen incompatibles entre sí. Como hemos visto, las dos primeras son la tesis de la Objetividad, la cual parece implicar al Cognitivismo, y el Internalismo. Como hemos visto, ambas, son tesis acerca del contenido de los JMs, es decir, son tesis semánticas. La tesis de la Objetividad involucra la idea de que los JMs poseen rasgos semánticos que hacen posible que los evaluemos como correctos o incorrectos. Si, como suele pensarse, esta tesis implica al Cognitivismo; entonces, un rasgo semántico de los JMs es el ser *apto para la verdad* (esto es lo que en inglés se denomina "Truth-Aptness"). Por su parte, el Internalismo involucra la idea de que los JMs son intrínsecamente motivacionales, lo cual sólo puede explicarse diciendo que ellos motivan acciones tan sólo en virtud de su contenido. Así, de acuerdo con el Internalismo, el contenido de un JM está constituido, al menos en parte, por la motivación para actuar de conformidad con él; en otras palabras, de acuerdo con esta tesis, los JMs posee un tipo de contenido que está *vinculado a la acción*. Pero, ¿por qué resulta tan problemático dar cuenta simultáneamente de estas dos tesis? ¿por qué nos parece tan dudosa la idea de comprometernos con una epistemología y una metafísica especiales, como las del Intuicionismo? La respuesta a estos interrogantes tiene que ver con la tercer tesis que, según Smith, genera el Problema Moral.

La mayor parte de los autores contemporáneos que trabajan en metaética consideran que la Psicología Humeana es la mejor teoría de que disponemos, para explicar la mente y el comportamiento humano. Por tal razón, según ellos, esta teoría debe constituir el marco de referencia de cualquier teoría que pretenda comprender la naturaleza de la moralidad. Ahora, de acuerdo con la Psicología Humeana, todo estado mental pertenece inevitablemente a una de dos clases: *creencias* o *deseos*. Las creencias son estados psicológicos que poseen contenido representacional. En la medida en que representan estados de cosas en el mundo, ellas pueden ser evaluadas como verdaderas o falsas, dependiendo de si los representan fielmente o no. En contraste, los deseos son estados que no poseen contenido representacional, en el mismo sentido en que lo poseen las creencias. Por tal razón, ellos no pueden ser evaluados en términos de verdad o falsedad. La característica distintiva de los deseos es que ellos poseen un vínculo directo con nuestras acciones; en otras palabras, ellos son estados psicológicos volitivos, que poseen fuerza motivacional, los cuales, en las condiciones apropiadas, pueden llegar a causar acciones en el mundo. La afirmación fundamental de la Psicología Humeana es las creencias y los deseos son *existencias diferentes*, en el sentido de que, para una creencia *C* y un deseo *D*, es perfectamente concebible que una persona posea *C* y a la vez carezca de *D*, y viceversa (Cf. Smith 1994, pp. 7-9). Algo que es importante notar, que rara vez se hace explícito en las discusiones en metaética, es que la Psicología Humeana, es el correlato psicológico de una tesis metafísica mucho más fundamental. Al respecto, John McDowell afirma:

If the world is, in itself, motivationally inert, and is also the proper province of cognitive equipment, it is inescapable that a strictly cognitive state –a conception of how things are, properly so called- cannot constitute the whole of a reason for acting. But the idea of the world as motivationally inert is not an independent hard datum. It is simply the metaphysical counterpart of the thesis that states of will and cognitive states are distinct existences [...]. (1998a, p. 83)

Según esto, sostener la Psicología Humeana implica comprometerse con una concepción, en la que el mundo está constituido por entidades que deben caracterizarse exclusivamente en términos de algo como las cualidades primarias de Locke, y en la que la subjetividad es la única fuente de motivación.

Tan pronto nos percatamos de que la Psicología Humeana está en el trasfondo de las discusiones en metaética contemporánea, podemos comenzar a entender por qué es tan problemática la idea de que una misma teoría puede dar cuenta de la Objetividad y del Internalismo. Por un parte, si queremos dar cuenta de la Objetividad, entonces debemos decir que los JMs son aptos para la verdad. A la luz de la Psicología Humeana, esto implica que debemos equipar a los JMs con creencias; en otros términos, debemos concebirlas como enunciados, cuya función primordial es *describir* a un mundo que es, en sí mismo, motivacionalmente inerte. Así, los JMs tendrán valor cognitivo; pero carecerán de fuerza motivacional. En consecuencia, por esta vía nos vemos obligados a rechazar al Internalismo. Por la otra parte, si queremos dar cuenta del Internalismo, entonces debemos decir que los JMs poseen un contenido que está vinculado a la acción. De acuerdo con la Psicología Humeana, esto quiere decir que debemos equipar a los JMs con deseos; o en otras palabras, que debemos concebir a los JMs, como enunciados, cuya función primordial es *expresar* los estados motivacionales del agente. De este modo, los JMs tendrán fuerza motivacional; pero no podrán ser evaluados como correctos e incorrectos. Por consiguiente, por esta vía nos vemos obligados a rechazar la Objetividad (Cf. Smith 1994, pp. 9-13). Como hemos dicho, la mayor parte de los teóricos en metaética dan sentado que la Psicología Humeana es verdadera. Por tal razón, se han movido entre estas dos posibilidades: o rechazar la Objetividad, o rechazar al Internalismo. La primera alternativa ha sido adoptada por Emotivistas clásicos, como A.J. Ayer y C.L. Stevenson, y por Teóricos del Error, como J.L. Mackie; a quienes les interesa dar cuenta, principalmente, de los aspectos subjetivos, motivacionales, del lenguaje moral. En la medida en que estos autores rechazan el Cognitivismo, sus posiciones pueden calificarse como "Antirrealistas"; puesto que su tesis fundamental es que el lenguaje moral no se refiere a objetos que se encuentran en la realidad material. En contraste, la segunda alternativa ha sido adoptada por Realistas Morales de todo tipo, incluyendo a los Naturalistas; cuyo interés principal es dar cuenta de la Objetividad. En virtud de ello, y dada su aceptación de la Psicología Humeana y de su correlato metafísico, estos autores deben realizar una defensa vehemente del Externalismo. Un intento muy interesante de dar cuenta de manera simultánea de la Objetividad y del Internalismo, puede encontrarse en las obras de R.M Hare y de Allan Gibbard. La estrategia de estos autores es cuestionar la idea de que, para explicar la Objetividad, necesariamente debemos comprometernos con el Cognitivismo. Según ellos, estados motivacionales como deseos poseen ciertas condiciones de corrección, las cuales pueden especificarse sin recurrir a la noción de verdad. Así, es posible pensar en los JMs como enunciados que expresan ciertos estados motivacionales del agente y que, aún así, pueden ser evaluados como correctos o incorrectos; dependiendo de la corrección o incorrección de esos estados. Desde esta perspectiva, entonces, se puede dar cuenta de la Objetividad y del Internalismo, sin comprometerse con el Cognitivismo. Hare denomina a su teoría Prescriptivismo; mientras que Gibbard la denomina Expresivismo Normativo. Sin embargo, es importante notar que estas teorías son formas de Antirrealismo, que se mueven dentro del marco de la Psicología Humeana. En ese sentido, la única innovación que podemos encontrar en ellas es cuestionamiento acerca de la relación entre la Objetividad y el Cognitivismo.

Creo que todas estas teorías son insatisfactorias, por una simple razón: dan por sentado que la Psicología Humeana, así como su correlato metafísico, son verdaderas. Una teoría metaética satisfactoria debe estar en capacidad de rechazar a la Psicología



Humeana; de modo que pueda dar cuenta tanto de la Objetividad como del Internalismo, sin rechazar el Cognitivism. Evidentemente, tal teoría debe ser muy próxima al Intuicionismo. Sin embargo, ella debe permitirnos concebir una forma nueva de conocimiento y una nueva visión del mundo; sin, por ello, llegar a sostener algo tan extravagante como la epistemología y la metafísica especiales que, en este capítulo, le hemos atribuido al Intuicionismo. El propósito del presente trabajo es sentar las bases para el desarrollo de una teoría metaética como la que acabamos de describir y aportar algunos argumentos que muestren que, efectivamente, ella es posible.

## 2. Los Inconvenientes del Antirrealismo Subjetivista

En el capítulo anterior vimos que las posiciones Realistas, incluido el Naturalismo, necesariamente deben comprometerse con la tesis controversial de que el Externalismo es verdadero. Por otra parte, el Intuicionismo, debido a la necesidad de dar cuenta del Internalismo, es llevado inevitablemente a la absurda posición, de sostener una epistemología y una metafísica especiales. El problema con estas dos posiciones es que, a pesar de sus diferencias, ambas se comprometen con la idea, derivada del Cognitivismo, de que el lenguaje moral se refiere a propiedades que hacen parte de la estructura de la realidad. En vista de ello, varios autores han sugerido que la posición más consistente, dentro de esta discusión, es el Antirrealismo. Esto quiere decir que una teoría metaética satisfactoria debe rechazar la pretensión de explicar la semántica del lenguaje moral, por referencia a propiedades reales (sean naturales o prescriptivas). Dado que, como veremos, a la base de esta posición se encuentra la Psicología Humeana, los autores que abogan por el Antirrealismo consideran que éste debe adoptar una forma subjetivista. De acuerdo con ellos, la semántica del lenguaje moral debe explicarse en términos de estados mentales que poseen fuerza motivacional; pues, como hemos dicho, la motivación es algo inherente a la subjetividad. En consecuencia, el rechazo, por parte de los adeptos a esta posición de todo compromiso metafísico; implica necesariamente la aceptación del Internalismo. Ciertamente, esto es algo a favor de esta posición. Sin embargo, aún cabe preguntarse ¿pueden las teorías Antirrealistas subjetivistas dar cuenta de la objetividad del lenguaje moral, aún cuando ellas rechazan el Cognitivismo? Si la respuesta a esta cuestión es afirmativa, entonces las teorías de este corte serán la mejor alternativa para explicar la semántica del lenguaje moral. Así, dado que en la literatura filosófica se encuentra teorías Antirrealistas subjetivistas, de las que se dice que, efectivamente, dan cuenta de la objetividad; debemos examinar cuidadosamente que tan satisfactorias puede llegar a ser. El propósito de este capítulo es realizar tal examen.

### 2.1. Dos Formas de Antirrealismo

John Mackie (1977) ha sostenido una forma de Antirrealismo subjetivista conocida como la Teoría del Error. Para sustentar su teoría, Mackie se vale de la analogía que hay entre el discurso acerca de colores y el lenguaje moral. Según él, el discurso acerca de colores es un ejemplo claro de en qué consiste la teoría del error. Al examinar cuidadosamente los enunciados en que usamos términos de color, él encuentra que ellos se comportan del mismo modo que los enunciados en los que empleamos términos para referirnos a algo como las cualidades primarias de Locke. En otras palabras, nuestro discurso acerca de colores tiene pretensiones de objetividad, esto es, tiene la pretensión de referirse a propiedades reales, capaces de existir con independencia de la experiencia que podamos tener de ellas. Sin embargo, la ciencia moderna ha mostrado que el color no es una

propiedad real; sino que corresponde, propiamente hablando, a lo que los empiristas denominaron cualidades secundarias. De acuerdo con la doctrina empirista, las cualidades secundarias tienen un carácter eminentemente subjetivo, pues dependen de la constitución interna del sujeto que las percibe. Con el color pasa algo similar: él es simplemente una percepción, la cual depende por completo del hecho de que poseemos un aparato sensorial determinado y no otro. Así, si el discurso acerca del color tiene la pretensión de referirse a propiedades reales; entonces, al no existir en el mundo tales propiedades, todos los enunciados de color serán sistemáticamente falsos. A partir de estas consideraciones, Mackie sostiene que el lenguaje moral adolece del mismo defecto: él tiene pretensiones de objetividad, tiene la pretensión de referirse a propiedades reales, capaces de existir con independencia de la experiencia humana. Sin embargo, dado el carácter motivacional inherente a nuestro lenguaje moral, para que sus pretensiones de objetividad sean satisfechas, es necesario comprometerse con una metafísica especial, así como con la epistemología correspondiente, análogas a las del Intuicionismo. Así, dada la implausibilidad de esta asunción metafísica, al igual que los enunciados de color, nuestros JMs son sistemáticamente falsos. Las propiedades morales, en consecuencia, deben concebirse más bien como análogas a las cualidades secundarias y, en ese sentido, deben ser vistas como algo puramente subjetivo. Ahora, debo confesar que no logro entender las motivaciones que llevan a Mackie a sostener una teoría como esta. Tal como lo veo, Crispin Wright comparte este desconcierto y por tal razón formula, de manera plenamente justificada, el siguiente interrogante: si nuestra práctica de hacer JMs es sistemáticamente errónea, ¿por qué los seres humanos continuamos envueltos en tal práctica? ¿por qué simplemente abandonamos una práctica que, desde ningún punto de vista, nos reporta conocimiento? (Cf. Wright 1992, pp. 11-2). Adicionalmente, es importante notar que la verdad de la Teoría del Error de Mackie depende completamente de que un análisis Realista del lenguaje moral (y del discurso acerca de colores) sea correcto. En los numerales 2.2 y 3.3 daremos argumentos que sugieren que un análisis con estas características es completamente erróneo. Por tal razón, podemos considerar a la Teoría del Error como un modelo Antirrealista bastante implausible.

Una forma de Antirrealismo mucho más plausible es el Expresivismo. Esta es una denominación que se emplea para referirse, en general, a la mayor parte de las posiciones No-Cognitivistas; debido a que un postulado fundamental de las mismas es que la función semántica de los JMs es *expresar* estados no cognitivos del sujeto, estados como deseos, emociones, sentimientos o actitudes de aprobación y desaprobación. Es importante distinguir a esta posición, algo en lo cual los Expresivistas insisten bastante, de ciertas posiciones Cognitivistas, las cuales pueden denominarse Realismo Subjetivista o Naturalismo Subjetivista. Según estas últimas, la función semántica de los JMs es *reportar* la posesión por parte del hablante de algún estado no cognitivo. Así, en el Realismo y el Naturalismo Subjetivistas, los JMs se conciben de tal modo que ellos pueden ser evaluados como verdaderos o falsos, en virtud de si el hablante posee efectivamente el estado no cognitivo que dice tener. En contraste, de acuerdo con el Expresivismo, la emisión de un JM debe ser vista como un acto de habla, cuya fuerza ilocucionaria difiere de la de una aserción, aproximándose mucho más a la de emisiones como «Aaaayyy», «Huurra!» o «Gooooool!».

Ahora, si bien los Expresivistas coinciden en concebir a los JMs como esencialmente vinculados a la acción, sus posiciones difieren al considerar la naturaleza de dicha conexión. Así, por ejemplo, A.J. Ayer y Allan Gibbard consideran que la conexión se da, en virtud de que los JMs expresan emociones que pueden tener un efecto causal sobre la acción (Cf. Ayer 1952; Gibbard 1990). En contraste, C.L. Stevenson y R.M. Hare consideran que la conexión se da, debido a que la principal función de los términos morales es recomendar u ordenar ciertas acciones. Según Stevenson, esto es posible

gracias a lo que él denomina *significado emotivo*. (Cf. Stevenson 1937; Hare 1950 y 1997).

Adicionalmente, los Expresivistas también difieren en la manera como abordan la cuestión de la objetividad del lenguaje moral. Por una parte, Ayer y Stevenson consideran que los JMs no son en sí mismos objetivos. La apariencia de objetividad se debe, según Ayer a que el uso de un JMs está determinado por una teoría moral. Dicha teoría está constituida por una serie de principios que determinan qué emociones deben experimentarse ante una clase definida de acciones. Así, las discusiones que se dan en ética no versan propiamente sobre la corrección o incorrección de un JM; sino, sobre la adecuación de la teoría ética que subyace el uso de tal Juicio. Los desacuerdos morales, entonces, deben ser vistos como el producto de un desacuerdo respecto a la teoría ética (Cf., Ayer 1952). De igual modo, Stevenson ha tratado de darle sentido a la noción de objetividad, indicando que los desacuerdos no sólo involucran creencias; sino que también puede haber desacuerdos respecto a actitudes e intereses. (Cf. Stevenson 1937).

Tal vez los intentos más interesantes de dar cuenta de la objetividad pueden encontrarse en la obra de Hare y Gibbard. Hare parte de la idea de que el significado de los JMs tiene dos componentes: uno prescriptivo y uno descriptivo. El componente descriptivo está ligado a los estándares que en una comunidad determinada se emplean para determinar qué acciones son, por ejemplo, buenas y qué acciones no. Gracias a él, los JMs pueden tener condiciones de verdad. Sin embargo, esto no les confiere objetividad; puesto que los estándares morales varían de una comunidad a otra. La objetividad, propiamente hablando, es dada por el elemento prescriptivo. Este elemento está ligado a la idea de que los JMs son esencialmente motivacionales, esto es, a la tesis del Internalismo. De acuerdo con Hare, esta caracteriza de los términos morales, la cual es independiente del contenido descriptivo, les confiere ciertos rasgos formales, que hacen posible que el vocabulario moral se rija por los principios de algo análogo a una lógica de los imperativos (Cf. Hare 1950 y 1997). En una dirección similar, Gibbard ha tratado de explicar la objetividad del lenguaje moral, mediante una noción no cognitiva de racionalidad. Según él, el hecho de que los JMs expresen estados que poseen fuerza motivacional, como emociones, no implica, como suele pensarse, que ellos no poseen condiciones de corrección. En efecto, es posible considerar a unas emociones como racionales y a otras como irracionales, dependiendo de qué tanto se ajustan ellas a ciertos principios de racionalidad. De acuerdo con estos principios, una emoción es racional si se produce ante la percepción de cierta clase de acciones; mientras que es irracional si se produce en una situación diferente a la que estipula el principio (Cf. Gibbard 1990).

## 2.2. La Crítica de McDowell al Expresivismo

I.

Como vimos, la formulación del Prescriptivismo de Hare y del Expresivismo Normativo de Gibbard incorpora dos elementos: una tesis semántica negativa (acerca de a qué clase de enunciados *no* pertenecen las adscripciones de valor) y una tesis metafísica negativa derivada de ella. Para este tipo de expresivistas, los juicios morales no son enunciados tales que su corrección consiste en ser *descripciones verdaderas* del mundo; de lo cual se deriva que los valores no son genuinas propiedades de las cosas. McDowell nota que asumir esta posición requiere estar en capacidad de dar una teoría positiva acerca de qué cuenta y qué no como una descripción verdadera del mundo. Las formulaciones del expresivismo suelen dar por sentada dicha teoría, en tanto que suponen una concepción del mundo que permite determinar qué enunciados, y qué enunciados no, son *true-aps*.

El propósito de McDowell es revelar los aspectos más importantes de dicha teoría y cuestionar la idea de que el expresivismo puede ser aceptado sobre la base de ella.

Según la concepción del mundo que está a la base del expresivismo, cómo las cosas realmente son es cómo son *en sí mismas*, esto es, *independientemente de cómo ellas se le presentan a los ocupantes de este o aquel punto de vista particular*. McDowell, de manera sugerente, propone que dando una interpretación literal de la noción de “punto de vista”, al observar las cosas debemos estar dispuestos a corregir nuestra perspectiva (i.e., estar dispuestos a mirar desde otros puntos de vista) cuando queremos determinar la *forma verdadera* de los objetos observados. Digo que esta afirmación se hace de manera sugerente, puesto que lleva a pensar que una correcta descripción de cómo las cosas son en sí mismas debe darse en términos de las propiedades que pueden adscribirseles con independencia de cualquier punto de vista particular, i.e., en términos su forma verdadera. El siguiente paso en la caracterización de McDowell es plenamente consecuente con esto: consiste en afirmar que tal concepción del mundo puede asociarse a la doctrina lockeana de que sólo las *cualidades primarias* son genuinos rasgos del mundo (en contraste con las cualidades secundarias).

¿Por qué, bajo esta concepción, las cualidades secundarias no pueden ser genuinos rasgos del mundo? Para McDowell no es suficiente decir que nuestro aparato perceptual no es universal, que sólo nosotros podemos percibir colores. Esto se debe a que tal afirmación todavía deja abierta la posibilidad de que nuestro aparato perceptual sea tal que nos permita percibir el mundo tal como es en sí mismo, mientras que el aparato perceptual de otros seres simplemente está lo suficientemente empobrecido para no permitirselos. El punto clave, según McDowell, consiste en reconocer que es posible dar una *explicación* de nuestra percepción del color únicamente en términos de cualidades primarias, lo que permite considerar a tal percepción como no verídica, pues no permite apreciar de manera clara e inmediata los verdaderos rasgos del mundo, aquellos que son responsables de producir tal percepción. Como hemos dicho, Mackie deriva su Teoría del Error precisamente de esta idea: nuestra percepción sistemáticamente nos lleva a pensar en las cualidades secundarias como genuinos rasgos del mundo, aun cuando no lo son.

Tomando en consideración esta concepción del mundo (de lo qué es un verdadera descripción del mundo y de lo qué es un enunciado *true-apt*), McDowell encuentra un paralelo entre la experiencia de color y la experiencia moral: así como aprendemos a hacer clasificaciones correctas de color (i.e., a usar *conceptos* de color) en vista de que nuestro aparato perceptual nos da ciertas experiencias; de igual modo aprendemos a hacer clasificaciones evaluativas correctas (y a *interesarnos* apropiadamente por las cosas que vemos bajo esas clasificaciones). El paralelo puede apreciarse más claramente en el siguiente pasaje:

In the argument about colors, we are led to appeal to the explanatory power of a description of the world in primary-quality terms, in order to exclude the suggestion that colour vision is a mode of awareness of something that is there anyway. The parallel suggestion, in the case of values, would be that the members of some specific set of values are genuine features of the world, which we are enabled to detect by virtue of our special affective and attitudinative propensities. (1998b, p. 200)

Sin embargo, McDowell nota que este paralelo puede verse amenazado por el hecho de que, en el caso de los valores, la sugerencia puede hacerse, valiéndose de algo respecto a lo cual no hay análogo en el caso del color; a saber, una filosofía de la mente (de corte humeano) según la cual hay una separación de género entre estados representacionales (capacidades cognitivas y su ejercicio) y estados motivacionales (los que los escritores del siglo XVIII denominaban pasiones o sentimientos). Tal filosofía amenaza la sugerencia,

puesto que al equiparar las percepciones de color y las experiencias morales se piensa en el ejercicio de nuestras disposiciones afectivas como en una expansión de nuestra percepción (como el ejercicio de una capacidad cognitiva); mientras que la filosofía del siglo XVIII excluye esto *a priori*.

Para McDowell una cuestión importante es hasta qué punto la adopción de tal filosofía de la mente realmente puede apoyar el expresivismo. Esta pregunta parece importante, puesto que existe la sensación de que la única razón por la cual se debe aceptar tal filosofía es la convicción en la verdad del no-cognitvismo (así, un argumento desde ella hacia este último tendría algo de circularidad. Cf. Smith, sobre la idea de que sólo la psicología humeana puede proveer una explicación teleológica de la acción). Por tal razón, McDowell pospone la discusión de dicha filosofía de la mente. En su lugar, se ocupa de tratar de especificar (y posteriormente cuestionar) cómo el expresivismo puede acuñar una explicación de la experiencia moral que corra en el mismo sentido que la explicación que se ofrece en el caso de las cualidades secundarias (esto es, una explicación a partir de una concepción del mundo como independiente de una perspectiva particular).

En la elaboración de tal explicación aparece la noción de “separación” (disentangling). Según el expresivismo, cuando se hace una adscripción de valor (cuando se hace un juicio moral), lo que en esta ocurriendo en ese momento puede separarse en dos componentes: por una parte, el ejercicio de una capacidad sensitiva que permite detectar un aspecto del mundo tal como es independientemente de una perspectiva particular; y por la otra, el ejercicio de una actitud no-cognitiva que permite enriquecer la concepción que se tiene de ese rasgo (esto es lo que constituye la perspectiva moral especial desde la cual se otorga valor al rasgo en cuestión). Como hemos visto, esta explicación es la que justamente encontramos en el Prescriptivismo de Hare y en el Expresivismo Normativo de Gibbard; y es la que les permite a estos autores responder al Problema Frege-Geach y dar cuenta de la noción de objetividad en el lenguaje moral. El propósito de McDowell es cuestionar la posibilidad de una explicación de la experiencia de valor que corra en estas líneas.

## II.

¿Para cada concepto de valor, es posible aislar un único rasgo del mundo (tal como es, independientemente de cualquier perspectiva particular) al cual respondan quienes usen competentemente tal concepto? Si la respuesta es afirmativa, esto implica que es posible para alguien que no pertenece a cierta comunidad moral adquirir competencia en el uso de sus conceptos morales, sin por ello adquirir su perspectiva moral particular. Tal competencia consistiría en reconocer cuál es el rasgo del mundo en virtud del cual se aplica el concepto moral a ciertas acciones (y así mismo estar en capacidad de predecir futuras aplicaciones del concepto), sin ejercer ni sentir ningún interés particular, sin tener la actitud no-cognitiva correspondiente (i.e., no sólo no compartir el punto de vista moral particular de la comunidad, sino ni siquiera tratar de darle sentido a, de entender, ese punto de vista particular). Según la caracterización del expresivismo que se vale de la “separación”, tal cosa es perfectamente posible, puesto que el ejercicio de una actitud no-cognitiva es un ingrediente extra, separable, en la competencia con el concepto moral. ¿Es posible un cuadro como este?

Como hemos dicho, Hare y Gibbard consideran que la explicación expresivista que McDowell está examinando puede sostenerse gracias a la importancia que en la metaética tiene la noción de superveniencia. Para Hare, las propiedades morales supervienen sobre propiedades no-morales de una clase aún por especificar. Según esto, el lenguaje moral satisface el siguiente requisito: siempre que se encuentren diferencias expresables en términos del nivel sobre el que las propiedades morales supervienen,

debe haber diferencias expresables en términos morales. A partir de ello, Hare deriva la tesis de la separación: es posible tener competencia en el uso de conceptos morales, sin que ello implique el ejercicio de las actitudes no-cognitivas que constituyen un punto de vista moral particular. Sin embargo, McDowell alega que de la satisfacción del requisito de la superveniencia no se sigue tal conclusión, puesto que de ello no se sigue que las propiedades no-morales sobre las que supervienen las propiedades morales constituyan un *género* definido. En palabras de McDowell:

[...] Supervenience leaves open this possibility [...]: however long a list we give of items to which a supervening terms applies, described in terms of the level supervened upon, there may be no way, expressible at the level supervened upon, of grouping just such items together. (*Idem*, p. 202)

Si esto es así (de hecho, pueden encontrarse ejemplos de ello en los diálogos socráticos de Platón), entonces no es posible que alguien tenga competencia en el uso de un término (en el sentido de poder predecir usos futuros del mismo) que funciona en el nivel sobre el que supervienen las propiedades morales y bajo el cual caigan todos los casos que, para alguien que es competente en el uso de un concepto moral, caen justamente bajo ese concepto moral.

El expresivista aún puede alegar que estas afirmaciones simplemente muestran que es *difícil* determinar cuál es el rasgo o propiedad no-moral sobre la que superviene una propiedad moral. Sin embargo, hay una serie de ideas acerca de la relación mente-mundo que permiten afirmar que sí *debe* existir tal rasgo. El propósito de McDowell, a continuación, es emplear las observaciones de Wittgenstein acerca del seguimiento de reglas para cuestionar esta maniobra (concretamente, para cuestionar las ideas acerca del significado y la relación mente-mundo que la soportan).

### III.

Este apartado es el núcleo de la argumentación de McDowell. Como se dijo arriba, el propósito aquí es cuestionar la idea de que debe existir un rasgo del mundo (tal como es en sí mismo, con independencia de cualquier perspectiva particular) al cual responde un hablante competente cuando emplea un concepto moral. Las reflexiones anteriores sugieren que no existe tal rasgo (que las clasificaciones morales se hace únicamente a partir de las propiedades morales que nos presenta nuestra experiencia evaluativa); sin embargo, el expresivista aún puede hacer un movimiento: afirmar que esto sólo muestra que es difícil especificar cuál es el rasgo no-moral al que estamos respondiendo (de ahí la dificultad para la teoría ética), pero que no logra mostrar que tal rasgo no existe. A esta tesis apuntan las consideraciones de McDowell acerca del seguimiento de reglas.

Las consideraciones acerca del seguimiento de reglas parten de la siguiente afirmación: hacer juicios sucesivos, en los cuales se aplica el mismo concepto a diferentes objetos (o acciones), pertenece a una práctica de seguir *haciendo la misma cosa* (going on *doing the same thing*). Parece correcto caracterizar tal práctica como una práctica guiada por reglas: un movimiento cuenta como estar haciendo la misma cosa, dentro de la práctica en cuestión, siempre y cuando se ajuste a las reglas. McDowell pone la cuestión en términos de exclusión: las reglas permiten determinar qué cuenta y que no como un movimiento en el que se está haciendo la misma cosa: "What count as doing the same thing, within the practice in question, is fixed by its rules" (*Idem*, p. 203). Un aspecto importante de las reglas es que son independientes de las respuestas y reacciones a las que propende alguien que ha aprendido la práctica misma. Esto quiere decir que es posible determinar, desde un punto de vista que es independiente de las respuestas que caracterizan a un participante en la práctica, si una serie de movimientos dentro de ella

cuentan como estar haciendo la misma cosa. En otras palabras, la corrección de los movimientos puede determinarse *objetivamente*, incluso sin involucrarse dentro de la práctica. La competencia dentro de la práctica de continuar haciendo la misma cosa requiere la existencia de dos elementos: 1) reglas independientes de las respuestas y reacciones de los participantes, y 2) un estado mental de *comprender* esas reglas. Un participante es competente cuando logra vincular estos elementos, cuando su comprensión de la regla le permite actuar de conformidad con ella, guiado por ella. McDowell expone esta idea acerca de la competencia en términos metafóricos:

The rules mark out rails along which correct activity within the practice must run. [...] Acquiring mastery of the practice is pictured as something like engaging mental wheels with this objectively existing rails. (*Idem*, p. 203)

Esta imagen puede adoptar dos formas. De un lado, las reglas puede formularse explícitamente, como una codificación de la práctica en términos independientemente accesibles (como un manual que puede llegar a conocerse sin necesidad de estar involucrado en la práctica). La competencia en la práctica puede concebirse como el *conocimiento* (posiblemente implícito) de lo que es expresado por tales formulaciones; a su vez, el comportamiento acertado, “correr a través de los rieles”, puede entenderse como un comportamiento dictado por pruebas de corrección (pruebas que dicen que tal o cual movimiento es correcto dentro de la práctica) cuyas premisas mayores son esas formulaciones (*Cf.* el modelo de explicación de la acción basado en silogismos prácticos o imperativos hipotéticos). La segunda alternativa se da en casos en que la práctica de aplicar conceptos se resiste a una codificación como ésta. En tal caso, la competencia en el uso de un concepto debe entenderse como la *captación de un universal*. Tal captación, así como el conocimiento de reglas estables, provee al participante en la práctica de la capacidad de comportarse correctamente, continuar haciendo la misma cosa o “correr a través de un riel que existe de manera independiente”.

McDowell quiere llegar a la siguiente conclusión:

[...] in fact the idea that the rules of a practice mark out rails traceable independently of the reactions of participants is suspect even in this apparently ideal case [i.e., “add 2”]; and if we insist that whenever there is going on in the same way there must be rules that can be conceived as marking out such independently traceable rails, we reveal a misconception of the sort of case in which correctness within a practice can be given the kind of demonstration we count as proof (*Idem*, p. 204)

En la visión que se está considerando de la aplicación de conceptos, los objetos (y sus rasgos) a que los conceptos se aplican son estructuras normativas estables (obsérvese la metáfora: “rieles que existen en el mundo independientemente y que guían la práctica”). La formulación de las reglas de la práctica debe decir que, dado tal o cual rasgo, debe aplicarse tal o cual concepto. La estabilidad de las reglas no es otra cosa que la estabilidad del mundo: en virtud de que el mundo es estable e independiente de cualquier perspectiva particular, las reglas que gobiernan la práctica de usar conceptos son estables e independientes de las respuestas y reacciones particulares de los participantes. Tener competencia en la práctica consiste, entonces, en tener una concepción del mundo (como independiente de cualquier perspectiva particular), tal que sus rasgos se aprecien como estructuras que guían la práctica de aplicar conceptos. Dicha concepción está constituida por reglas explícitas que codifican la práctica. Dado esto, como dice McDowell, todo movimiento correcto dentro de la práctica puede



entenderse como una acción dictada por una prueba o deducción en la que la formulación de una regla funciona como premisa mayor<sup>1</sup>:

Para todo r, aplíquese (el concepto) “R”  
Este x es un r  
∴ Este x es “R”

Esta visión es la base para el expresivismo, puesto que permite separar el ejercicio de capacidades cognitivas (las cuales se asocian a este procedimiento guiado por reglas) del ejercicio de la actitud no-cognitiva que constituye un punto de vista moral particular. Por tal razón, McDowell quiere rechazarla, para negar la posibilidad de separación. Ahora, nótese que esta visión, en tanto que da cuenta de la aplicación de conceptos, de la relación mente-mundo, incorpora dos elementos: uno metafísico, otro psicológico. La estrategia de McDowell consiste en concentrarse en el elemento psicológico, para mostrar su implausibilidad (para mostrar que no es posible que exista un estado consistente en *comprender (explícita o implícitamente) una regla* que es independiente de las respuestas y reacciones de los participantes en la práctica).

Según la imagen, la comprensión de una regla como “sumar 2” es un mecanismo psicológico que causa (churns out) el comportamiento adecuado (en la práctica de continuar la serie 2, 4, 6, 8, ...) con la precisión que tiene, p.e., el mecanismo de un reloj. Si alguien continúa correctamente la serie y podemos suponer que lo hace debido a que comprende la regla y la está siguiendo, entonces se puede formular la *hipótesis* de que el estado mental de comprender la regla subyace a su comportamiento. El problema, siguiendo a Wittgenstein, es tratar de especificar qué es aquello que manifiesta la comprensión de la regla. Toda manifestación de comprensión debe ir acompañada por cierta *evidencia*, i.e., por una serie finita (un parte de la serie potencialmente infinita) de movimientos que aparentemente son causados por el estado de comprender la regla. En tanto que tal evidencia es finita, su presencia es plenamente compatible con la idea de que es posible que en otra ocasión el comportamiento difiera de lo que es considerado un comportamiento correcto y no por simple error: es posible que al llegar a 1000 continúe la serie con 1004, 1008, ... En caso de que esto ocurra y la persona no se comporte de este modo por error, entonces debe decirse que la hipótesis de que ella estaba actuando debido a su comprensión de la regla “sumar 2” era errónea. El problema es que, para toda práctica, existe esta posibilidad. Así, la atribución de un estado de comprender una regla siempre trasciende la evidencia disponible, la cual permite justamente tal atribución; por lo cual parece vacío recurrir a tales estados para caracterizar el desempeño en una práctica.

McDowell hace dos aclaraciones. Primera, el problema no es una forma de *escepticismo acerca de las otras mentes*, que no aplica en el caso propio. No es posible decir “yo sé que estoy siguiendo la regla ‘sumar 2’”, porque de todos modos queda abierta la posibilidad de que alguien, al llegar a 1000, continúe la serie 1004, 1008, ... sin sentir que está actuando a la deriva (i.e., sintiendo que su comportamiento se ajusta a una regla). Segunda, es erróneo pensar en el problema como la siguiente forma de

---

<sup>1</sup> Sobre esto, ver las críticas de McDowell a la visión ética inspirada en Aristóteles, según la cual una concepción de la felicidad funciona como premisa mayor en las deducciones que llevan a realizar acciones particulares moralmente correctas. Las críticas de McDowell a este modelo se aplican de igual modo allí (de hecho McDowell también se vale allí de las consideraciones sobre el seguimiento de reglas). La conclusión que permite derivar “Virtue and Reason” (en: 1998) es que la concepción de la felicidad no es dada conceptualmente, y por tanto, no puede llevar a la acción moral del modo caracterizado en este modelo. La felicidad es dada no-conceptualmente!

escepticismo: no podemos saber que el comportamiento de otros no será azaroso (no se ajustará a reglas). Para McDowell, lo importante del argumento no es dejar abierta la posibilidad de que alguien se comporte como en el caso de quien continua la serie 1004, 1008, ... De hecho, tenemos confianza en que tales casos no ocurrirán. La cuestión es poner en duda la idea de que tal confianza se basa en el hecho de que el comportamiento es guiado por ciertas estructuras fundamentales (reglas estables independientes de las respuestas de los participantes en la práctica dada) a través del mecanismo psicológico descrito (el estado de comprender esas reglas): “[...] the argument aims, not to undermine this confidence, but to change our conception of its grounds and nature” (*Idem*, p. 205). La pregunta es, entonces, ¿cuál será la base de esa confianza?

Estas consideraciones tienen como consecuencia el que es imposible, para los seres humanos, captar cualquier rasgo del mundo, con independencia de nuestra perspectiva: el estado mental correspondiente a captar un rasgo del mundo de este modo sería análogo al estado mental consistente en captar una regla estable independiente de las respuestas de los participantes en la práctica dada. De esto se sigue, igualmente, que la imagen del mundo que subyace a la idea de que es posible explicar el comportamiento de un participante en una práctica, mediante la atribución del estado mental de comprender una regla estable e independiente, es una imagen bastante dudosa. Es decir que, el elemento metafísico que está a la base de tal visión es igualmente dudoso. De modo que, difícilmente, podemos pensar en el mundo como en un dominio constituido por estructuras normativas estables e independientes de nuestras prácticas y actitudes evaluativas (estructuras como “rieles que existen en el mundo independientemente y que guían la práctica”). Con lo cual, los supuestos básicos del expresivismo, los cuales hacen posible la tesis de la separación, resultan insostenibles.

El expresivismo, en consecuencia, no puede dar cuenta de una noción de objetividad, del modo en que pretenden Hare y Gibbard, pues al considerar que los términos morales simplemente tienen una función expresiva, deben comprometerse con una visión del mundo que, como McDowell muestra, es insostenible.

### 3. Objetividad, Antirrealismo y Experiencia Intersubjetiva

En el capítulo anterior señalamos los problemas que presentan las teorías metaéticas de corte antirrealista-subjetivista. Dijimos que, algunas de ellas no pueden dar cuenta de la objetividad del lenguaje moral; mientras que otras, al intentar hacerlo, se comprometen con una visión del mundo bastante problemática. Ante esta situación, el panorama parece un poco desalentador. Por una parte, tenemos teorías realistas-naturalistas, las cuales pueden explicar la objetividad; pero niegan el Internalismo. Por otra, tenemos teorías antirrealistas-subjetivistas, que dan cuenta de los aspectos motivacionales del lenguaje moral (pues lo relacionan con emociones u otros estados mentales que poseen fuerza motivacional); pero son incapaces de explicar la objetividad. Y finalmente, tenemos teorías Intuicionistas, que pueden dar cuenta del Internalismo y de la objetividad; pero al costo de tener que comprometerse con una metafísica y una epistemología especiales. ¿Qué hacer frente a este cuadro? ¿debemos comenzar a pensar que es imposible dar una explicación satisfactoria de nuestro lenguaje moral? En este capítulo, quiero sugerir que sí es posible dar una explicación satisfactoria del lenguaje moral. Tal explicación puede darse, a partir de una teoría que en muchos aspectos se aproxima al Intuicionismo, pero que no adolece de los problemas que hemos encontrado en él. El propósito principal de este capítulo es esbozar dicha teoría.

Para cumplir con este propósito, creo que es necesario comenzar por considerar dos preguntas clave. La primera es: ¿qué requisitos deben satisfacerse para que podamos decir que el lenguaje moral es objetivo? La segunda es: ¿necesariamente toda forma de antirrealismo moral debe ser subjetivista? Al considerar cuidadosamente estas preguntas, es posible comenzar a entrever qué tipo de teoría necesitamos para explicar el lenguaje moral.

#### 3.1. El Problema de la Objetividad del Lenguaje Moral

Comenzaremos con la cuestión acerca de la objetividad. De manera intuitiva, podemos decir que la idea de que el lenguaje moral es objetivo involucra algo como lo siguiente: decir de un JM que es objetivo es concebirlo como una norma de conducta universal, esto es, como una norma que no está determinada por las creencias e intereses particulares de un agente ni por las prácticas socioculturales propias de su contexto particular.

Como hemos dicho, el problema de determinar si el lenguaje moral es objetivo es fundamentalmente un problema de carácter *semántico*. En efecto, para atribuirle objetividad a los JMs, es indispensable que ellos posean unos rasgos semánticos específicos que hagan posible tal atribución. Podremos determinar con precisión si los JMs son objetivos o no, sólo si de antemano hemos mostrado cuál es la correcta especificación de su significado y hemos investigado si tal especificación revela que ellos poseen los rasgos semánticos necesarios para atribuirles objetividad. Por esta razón, la discusión en metaética acerca de la objetividad ha estado ampliamente relacionada con la

cuestión de si la clave para especificar el significado de los JMs es la noción de verdad, o alternativamente, la noción de ser apto para la verdad (lo que en inglés se denomina "Truth-Aptness"). En vista de esto, la taxonomía de las posiciones que han surgido en torno al problema debe trazarse teniendo en cuenta dos criterios, a saber, cómo cada una de las teorías en cuestión responde a dos preguntas cruciales:

- i. ¿Los JMs son objetivos?
- ii. ¿Los JMs son aptos para la verdad?

La pregunta (ii) constituye el corazón de la disputa entre Cognitivistas y No-Cognitivistas. Como sabemos, los Cognitivistas son aquellos que piensan que los JMs tienen un *valor cognitivo*. Esto quiere decir que, para ellos, cuando las cosas funcionan bien, un JM nos da conocimiento acerca del mundo, acerca de la sociedad o acerca de la psicología de una persona específica. Dado el estrecho vínculo que hay entre conocimiento y verdad, lo que define al Cognitivismo es la idea de que los JMs son aptos para la verdad; así, el Cognitivismo responde afirmativamente a (ii). En contraste, el No-Cognitivismo niega que los JMs tienen valor cognitivo. Según él, la característica esencial de los JMs es que tienen una conexión intrínseca con la acción. Por tal razón, su función primaria es *expresar* estados mentales que desempeñan un rol en la producción de nuestras acciones, v.g., estados como emociones, sentimientos o actitudes de aprobación y desaprobación. Por esta razón, el No-Cognitivismo responde negativamente a (ii).

Ahora, la posición tradicional con respecto al problema consiste en pensar que ser Cognitivista implica necesariamente ser objetivista acerca del JM. Esta posición puede resumirse en la siguiente afirmación: que los enunciados de un discurso sean aptos para la verdad es una *condición necesaria y suficiente* para que ellos sean objetivos. Desde este punto de vista, entonces, las respuestas a (i) y a (ii) son afirmativas. Esta posición es sostenida típicamente por los Intuicionistas y por los Realistas Morales de cualquier variedad (p.e., Railton, 1986; Sturgeon, 1988)<sup>2</sup>. Sin embargo, algunos autores aceptan el Cognitivismo, pero rechazan la tesis de que los JMs son objetivos. Ellos responden negativamente a (i) y afirmativamente a (ii). Esta tesis puede ser sostenida de dos maneras distintas. En primer lugar, puede ser sostenida en la forma de una Teoría del Error, como la propuesta por Mackie. Como vimos, Mackie acepta una parte de la posición tradicional, a saber, que los enunciados de un discurso sean aptos para la verdad es una *condición necesaria* para que ellos sean objetivos; sin embargo, rechaza que esto sea una condición suficiente, ya que es posible que las entidades acerca de las cuales hablan esos enunciados no hagan parte de la estructura de la realidad. Si tal cosa ocurre, entonces todos los enunciados de ese discurso serán falsos y no tendrá sentido decir de ellos que son objetivos. Según Mackie, esto es justamente lo que ocurre con los JMs: ellos son aptos para la verdad, pero hablan acerca de propiedades prescriptivas, propiedades que guían la acción, las cuales no se encuentran en la realidad. Todos los JMs son falsos, y en consecuencia, no tiene sentido decir de ellos que son objetivos. En segundo lugar, esta posición puede ser sostenida en la forma de un Relativismo como el propuesto por Gilbert Harman (2000). Según él, si bien los JMs son aptos para la verdad, las condiciones de verdad de un JM específico *son relativas* a un acuerdo tácito entre las personas que conforman la comunidad a la cual pertenece el hablante que emitió ese JM.

---

<sup>2</sup> Si bien las variedades de Realismo Moral surgen del desacuerdo acerca de las *particularidades metafísicas de las propiedades morales*, la afirmación de que el significado de los JMs está constituido de manera primordial por sus condiciones de verdad y de que, en virtud de ello, los JMs son objetivos puede ser vista como la tesis central de cualquier forma de Realismo Moral.

Un mismo JM, entonces, podrá ser verdadero en relación a una comunidad y falso en relación a otra; lo cual, como es evidente, quiere decir que no es objetivo.

Por otra parte, entre los No-Cognitivistas hay quienes piensan que los JMs sí son objetivos, que responden afirmativamente a (i), y quienes aparentemente piensan que no lo son, que responden negativamente a (i). Los No-Cognitivistas que responden afirmativamente a (i) deben comprometerse con la idea de que, contrario al pensamiento tradicional, el que los enunciados de un discurso sean aptos para la verdad *no es condición necesaria ni suficiente* para que tales enunciados sean objetivos. Tales No-Cognitivistas, entonces, deben comenzar por especificar qué es lo que, según ellos, es constitutivo de la noción de objetividad. Como vimos, R.M. Hare y Allan Gibbard han dado pasos significativos en esta dirección. Según ellos, los JMs responden a una normatividad distinta a la verdad. Hare (1952 y 1997) ha propuesto el Prescriptivismo, en el cual se afirma que los JMs, junto con los imperativos, pertenecen a una clase de expresiones prescriptivas. Teniendo en cuenta esto, él explica la objetividad de los JMs en términos de su ajuste a los principios de algo análogo a una *lógica de los imperativos*. Gibbard (1990), por su parte, explica la objetividad mediante una noción no-cognitivistista de *racionalidad*. Según él, las emociones poseen condiciones de racionalidad: hay emociones racionales y emociones irracionales. De este modo, la corrección de un JM está determinada por la racionalidad de la emoción que él expresa: un JM es correcto cuando expresa una emoción racional, e incorrecto cuando expresa una emoción irracional. Como sabemos, Gibbard denomina a esta teoría Expresivismo Normativo.

Finalmente, algunos autores han sugerido que hay formas de No-Cognitivismo en las que se rechaza por completo la objetividad (Cf. Hare 1997, pp. 115-38). Esta tesis suele atribuirse a la versión clásica del No-Cognitivismo, a saber, el Emotivismo de A.J. Ayer (1952) y de C.L. Stevenson (1937 y 1944). En esta posición, al igual que en la de Hare y Gibbard, un JM simplemente da expresión a una emoción que actualmente experimenta la persona que lo emite. Sin embargo, aquí no hay ningún criterio que permita evaluar tales emociones; por lo cual, no es posible criticar los JMs morales de ninguna persona. Una consecuencia de esto es que la moralidad no puede desempeñar el rol que usualmente le atribuimos, a saber, regular el comportamiento de las personas dentro de una sociedad.

Podemos resumir lo dicho hasta acá, mediante el siguiente cuadro:

|                       | <b>Cognitivismo</b>             | <b>No-Cognitivismo</b>                    |
|-----------------------|---------------------------------|---|
| <b>Objetivismo</b>    | Intuicionismo<br>Realismo Moral | Prescriptivismo<br>Expresivismo Normativo |
| <b>No-Objetivismo</b> | Teoría del Error<br>Relativismo | Emotivismo Clásico                        |

## 3.2. Criterios de Objetividad

Creo que, en lo fundamental, la posición tradicional es correcta: concebir a los JMs como aptos para la verdad es una *condición necesaria y suficiente* para atribuirles objetividad. Sin embargo, no estoy seguro de que esto implique que, para dar cuenta de la objetividad, necesariamente debemos comprometernos o con el Intuicionismo o con alguna forma de Realismo Moral. El propósito de esta sección es tratar de justificar esta afirmación. Para ello, debemos indagar un poco más acerca de qué está involucrado en la noción de objetividad.

Al inicio de la sección anterior dijimos que un JM es objetivo sólo si es una norma de conducta universal, esto es, una norma que no está determinada por las creencias e intereses particulares de un agente ni por las prácticas socioculturales propias de su contexto específico. Como indicamos, esta es apenas una caracterización intuitiva de la noción de objetividad, una caracterización que aún debemos refinar un poco más. ¿Qué está involucrado, entonces, en la noción de objetividad? Creo que, cuando decimos de los JMs que son objetivos, debemos comprometernos con dos afirmaciones importantes acerca de ellos. Si pensamos que un JM es objetivo sólo si es una norma de conducta universal, entonces podemos trazar una distinción entre los JMs que efectivamente son objetivos y los JMs que una persona está dispuesta a admitir. No necesariamente todos los JMs que una persona está dispuesta a admitir corresponden a normas de conducta universales: algunos de sus JMs pueden derivarse de consideraciones puramente particulares o pueden estar orientados simplemente a la satisfacción de intereses egoístas. Las dos afirmaciones constitutivas de la noción de objetividad se derivan de esta distinción. En primer lugar, trazar la distinción implica que hay situaciones en las que el JM de una persona no corresponde con lo que universalmente debería hacerse en esa situación. En tal caso, el JM de esa persona es incorrecto, puesto que no es objetivo. Así, hay JM correctos e incorrectos: son correctos cuando corresponde con lo que universalmente debe hacerse e incorrectos cuando no corresponden con lo que universalmente debe hacerse. De este modo, la primera afirmación es la siguiente:

(1) Decir de los JMs que son objetivos implica necesariamente atribuirles ciertas *condiciones de corrección*.

Tal como lo veo, esto nos lleva a sostener una tesis propiamente *semántica*: el significado de los JMs es tal que ellos son aptos para ser evaluados como correctos o incorrectos. A primera vista, entonces, el análisis No-Cognitivista es erróneo, puesto que no es claro cómo expresiones lingüísticas cuya función primaria es expresar estados mentales no-cognitivos, como emociones, pueden ser susceptibles a este tipo de evaluación. En segundo lugar, si evaluamos como correctos a los JMs de una persona que corresponden a JMs objetivos, entonces los criterios que permiten determinar la corrección de un JM no pueden ser subjetivos. Es evidente que, si tales criterios fueran subjetivos, los JMs que en virtud de ellos se evalúan como correctos serían correctos únicamente para el individuo o para el grupo de individuos que aceptara tales criterios de corrección (tal como ocurre, por ejemplo, en el Relativismo de Harman) y no de manera universal. Así, la segunda afirmación que constituye la noción de objetividad es la siguiente:

(2) Los JMs son objetivos, si sus condiciones de corrección son *independientes* del pensamiento particular de los individuos que emiten esos JMs; o en otras palabras, si sus condiciones de corrección son *externas* al discurso mismo.

Si los JMs han de ser objetivos, ellos deben ser *responsables* frente a una instancia externa al pensamiento particular de los individuos que los emiten, externa al discurso mismo. Creo que (1) y (2) son lo único esencial a la noción de objetividad. Se puede decir, entonces, que (1) y (2) expresan las condiciones necesarias y suficientes para la objetividad.

Algunos autores, sin embargo, piensan que la siguiente afirmación es la que mejor caracteriza en qué consiste la objetividad:

(3) los JMs son objetivos si y sólo si en todo *desacuerdo moral* necesariamente una de las partes está cometiendo un error (Brink, 1989; Wright, 1992).

Según ellos, en (3) se presenta una condición necesaria y suficiente para la objetividad. Nótese que (3) implica que para sustentar la idea de que un discurso es objetivo, necesariamente debemos comprometernos con la idea de que los criterios de corrección de los JMs son lo suficientemente determinados como para dejar de lado cualquier posible indecisión en un desacuerdo. Cuando hay indecisión, ella se debe a que carecemos de información suficiente o nuestras capacidades cognitivas están funcionando erróneamente. Dummett ha denominado “Principio de Bivalencia” a la tesis de que, en un discurso, de entre dos enunciados contrarios necesariamente uno ha de ser verdadero y el otro falso, con independencia de que los hablantes tengan la capacidad de llegar a conocer cuál es verdadero y cuál falso. Así, parece que detrás de (3) está presente la idea de que un discurso cuyos enunciados son objetivos necesariamente es un discurso en el cual tiene plena validez algo análogo al Principio de Bivalencia. En ese sentido, (3) es un requisito mucho más fuerte que la conjunción de (1) y (2); pues parece posible satisfacer estos últimos sin, por ello, tener que comprometerse con el Principio de Bivalencia. En vista de esto, debemos considerar los siguientes interrogantes: ¿para que el lenguaje moral sea objetivo, es necesario que él que satisfaga (3)? ¿es posible desarrollar una teoría en la que el lenguaje moral se caracterice, de tal modo que sea objetivo tan sólo por satisfacer (1) y (2)? (Cf. Shafer-Ladau 1994, p. 333)

Si la única manera de dar cuenta de la objetividad del lenguaje moral es mostrándolo como un discurso en el que se satisface (3), entonces las únicas teorías metaéticas capaces de dar cuenta de la objetividad son el Intuicionismo y las diversas formas de Realismo Moral. Sin embargo, como hemos visto, hay razones, independientes a consideraciones acerca de la objetividad, por las que estas teorías son problemáticas. Por otra parte, parece que lo que hay detrás del Prescriptivismo de Hare y del Expresivismo Normativo de Gibbard, es un intento por caracterizar el lenguaje moral, de un modo tal que él sea objetivo tan sólo por satisfacer (1) y (2). Sin embargo, los argumentos de McDowell, han mostrado que estas teorías son igualmente problemáticas. A luz de estas consideraciones, podemos decir que una teoría satisfactoria acerca del lenguaje moral debe retener aspectos importantes de estas dos clases de teorías. En primer lugar, ella no debe renunciar a la idea de que una explicación de la objetividad requiere de la atribución de aptitud para la verdad a los JMs. Esto es precisamente en lo que se equivocan el Prescriptivismo y el Expresivismo Normativo; pues, al tratar de especificar condiciones de corrección que no tienen que ver con el conocimiento del mundo, esto es, condiciones de corrección No-Cognitivas, deben llevar a cabo un sospechoso movimiento de separación. Pero, en segundo lugar, una teoría satisfactoria debe ofrecer una explicación de la objetividad en la que no se comprometa con la verdad del Principio de Bivalencia; pues la aceptación de dicho principio lleva, o bien a concebir un dominio de propiedades prescriptivas, como ocurre en el Intuicionismo; o bien a explicar la verdad de los JMs por referencia a propiedades naturales y a negar el Internalismo, como ocurre en las diversas variedades de Realismo Moral. ¿Cómo podemos desarrollar este tipo de teoría? ¿cómo

podemos atribuirle aptitud para la verdad a los JMs, sin por ello tener que comprometernos con el Principio de Bivalencia? Para tratar de responder a esta cuestión, es necesario abordar de una vez por todas la segunda pregunta clave que señalé al inicio de este capítulo, a saber: ¿necesariamente toda forma de Antirrealismo Moral debe ser subjetivista?

### 3.3. El Modelo Antirrealista de Dummett

En varios escritos, Michael Dummett ha propuesto un modelo general para el Antirrealismo, que puede aplicarse a diferentes tipos de discursos. Si bien Dummett desarrolla este modelo principalmente para explicar la semántica del lenguaje matemático, de todas formas, considero que él es de interés para la cuestión que ahora nos ocupa; pues, a pesar de ser antirrealista, este modelo permite atribuirle objetividad a cualquier discurso que se caracterice de acuerdo con él.

Para trazar la distinción entre una semántica Realista y una Antirrealista, Dummett parte de una serie de consideraciones acerca de en qué consiste comprender el significado de un enunciado. Según él, un discurso Realista posee lo que en filosofía del lenguaje se denomina una semántica Veritativo-Condiciona. Según esto, en un discurso Realista, la explicación de en qué consiste comprender el significado de un enunciado se da en términos de la noción de *condiciones de verdad*: comprender un enunciado es captar, aprehender, sus condiciones de verdad. Un aspecto importante de un discurso Realista es que, en él, las condiciones de verdad que se le atribuyen a los enunciados son concebidas como *trascendentes*, esto es, como condiciones que dependen de aspectos de la realidad, los cuales existen de manera determinada e independiente del conocimiento que los hablantes puedan tener de ellos. Esto implica que, según esta concepción, el valor de verdad de un enunciado depende única y exclusivamente de cómo es la realidad; la realidad es el criterio último en virtud del cual se evalúa a cada enunciado. A partir de esto, Dummett afirma que una característica importante de un discurso cuya semántica es Realista, es que la lógica de sus enunciados respeta plenamente el Principio de Bivalencia, esto es, el principio según el cual, para cualquier enunciado y su negación, necesariamente uno de ellos es verdadero y el otro falso. Esta es la razón por la que, en el numeral anterior, dijimos que una teoría que pretenda explicar la objetividad del lenguaje moral tratando de satisfacer (3), necesariamente debe comprometerse con el Principio de Bivalencia y con, o bien el Intuicionismo, o bien alguna forma de Realismo Moral. En contraste, según Dummett, en un discurso Antirrealista la explicación de en qué consiste comprender el significado de un enunciado se da en términos de algo similar a la noción wittgensteiniana de *uso*: comprender un enunciado es saber cómo usarlo en situaciones de habla específicas. Cómo debemos usar un enunciado es algo que está determinado por ciertas *condiciones de uso*, o alternativamente, por ciertas *condiciones de asertabilidad*, las cuales *justifican* el uso del enunciado en una situación específica. De tal modo, en este tipo de discurso, la corrección o incorrección de un enunciado se determina, en virtud de si es justificado usarlo o no en determinadas circunstancias. Por tal razón, algo esencial a un discurso Antirrealista es la posesión, por parte del hablante, de un *método objetivo de justificación* para sus preferencias lingüísticas. Precisamente, debido a esto, Dummett considera que una teoría Antirrealista de este tipo es lo que puede denominarse una Teoría Justificacionista del significado. Algo que es importante resaltar es que lo anterior no implica un rechazo, por parte del Antirrealismo, de la noción de verdad. De hecho, Dummett considera que esta noción es esencial para comprender la cognición humana. Por tal razón, las afirmaciones anteriores deben entenderse, más bien, en el siguiente sentido: en un discurso Antirrealista no se prescinde de la noción de verdad; sino que se



le interpreta en términos de la noción de condiciones de justificación. Así, para el Antirrealismo decir de un enunciado que es verdadero es equivalente a decir que es justificado usarlo en determinadas circunstancias (Cf. Dummett 1978a; 1991, pp. 322-51; 1993a; 2006, pp. 56-72).

Hay dos aspectos importantes de esta doctrina que debemos tener en cuenta. El primero de ellos es que, de acuerdo con ella, en un discurso Antirrealista no se respeta el Principio de Bivalencia. Para apreciar esto, consideremos, por ejemplo, un enunciado matemático  $A$ . Si no disponemos de una demostración de  $A$ , o de una demostración de  $\neg A$ ; parece que no hay nada que nos permita determinar con exactitud si  $A$  es verdadero o falso. Es posible que nadie esté en capacidad de construir alguna de estas pruebas o que, de hecho, sea imposible construirlas; en tal caso, el valor de verdad de  $A$  seguirá siendo algo incierto. Pero, incluso, si alguien está en capacidad de construir alguna de estas pruebas o si es posible construirlas, aún cuando actualmente nadie pueda hacerlo; de todas formas, el valor de verdad de  $A$  seguirá siendo incierto, mientras no dispongamos de alguna de ellas, esto es, mientras no haya algo que justifique la aserción de  $A$ .

El segundo aspecto importante de esta doctrina es que ella no es una forma de subjetivismo. En efecto, como hemos dicho, un discurso Antirrealista dispone de un método objetivo que permite justificar el uso de un enunciado en circunstancias específicas. Así, por ejemplo, en el discurso matemático se pueden construir pruebas para justificar una aserción; en el discurso acerca de colores se pueden recurrir a la percepción visual para justificar una aserción; etc. Detrás de la afirmación de que estos métodos de justificación son objetivos, está la idea de que ellos se rigen por una normatividad que, en buena medida, es independiente de nuestros intereses particulares. Consideremos, por ejemplo, las siguientes afirmaciones de Dummett acerca del lenguaje matemático:

[...] alguien que acepte el punto de vista intuicionista en las matemáticas no puede inclinarse a adoptar la imagen platonista. ¿Debe irse entonces al siguiente extremo opuesto y aceptar la imagen de que creamos las matemáticas sobre la marcha? Adoptar esta imagen implica pensar con Wittgenstein que en matemáticas somos *libres* en todo momento; para ningún paso que demos hemos sido forzados por una necesidad externa a nosotros, sino que hemos elegido libremente. Esta imagen no es la única alternativa. Si pensamos que los resultados matemáticos se nos imponen en algún sentido desde afuera podríamos tener más bien la imagen de una realidad matemática no existente sino hasta que comienza a existir cuando la pruebo. Nuestras investigaciones le dan existencia a lo que no la tenía, pero no es a nuestra creación a lo que dan existencia.

[...] Esto muestra cómo es posible sostener que la sustitución intuicionista de una explicación del uso de un enunciado por la explicación de sus condiciones de verdad como la forma general de explicación del significado debería aplicarse a todos los campos del discurso sin pensar por ello que creamos el mundo; podemos abandonar el realismo sin caer en un idealismo subjetivista. (1978a, p. 84)

No podemos construir cualquier prueba matemática que queramos, ni podemos construir una prueba matemática del modo en que queramos. Del mismo modo, no podemos tener cualquier experiencia de color que queramos, ni podemos tener experiencia de color bajo las condiciones que queramos. Las pruebas matemáticas y las experiencias de color dependen, por igual, de ciertas restricciones externas y de ciertos rasgos subjetivos inherentes a la condición humana, como ser racional o poseer un aparato perceptual determinado. En ese sentido, los objetos matemáticos y los colores cobran existencia, a partir de la interacción entre rasgos objetivos y rasgos subjetivos universalmente compartidos. Debido a esto, Dummett considera que no hay ningún problema en suscribir

la tesis de que los enunciados de un discurso Antirrealista son *responsables frente a una realidad externa* al sujeto que emite tales enunciados. Sin embargo, dada la negación del Principio de Bivalencia, el Antirrealismo debe concebir a esta realidad como algo *indeterminado*; como algo que, en ciertas ocasiones, es incapaz de hacer verdaderos o falsos a algunos de nuestros enunciados. Para comprender esto, consideremos el discurso acerca de colores. En efecto, como hemos dicho, no hay ningún problema en decir que nuestros predicados de color se refieren a rasgos que hacen parte de una realidad externa, la cual cobra existencia gracias a nuestra experiencia de color. Pero esta realidad es indeterminada; puesto que, por ejemplo, con respecto a una muestra de color que se encuentra exactamente en el punto medio entre el rojo y el naranja, bajo ninguna circunstancia, nuestra experiencia nos permite decir si ella es roja o naranja. Aún cuando disponemos de un método de justificación, esto es, aún si tenemos una experiencia de tal muestra, no podemos justificar la aserción del enunciado «ella es roja» ni la aserción del enunciado «ella es naranja» (Cf. Dummett 1978a, 1978b).

A la luz de esta caracterización, ahora debemos considerar la cuestión si este modelo Antirrealista es adecuado para explicar la semántica del lenguaje moral. Al respecto, Crispin Wright ha cuestionado la idea de que podamos emplear los planteamientos de Dummett para abordar las cuestiones metaéticas. A su modo de ver, el problema con el Antirrealismo de Dummett es que él es compatible con la idea de que los enunciados de un discurso son responsables frente a una realidad externa al hablante. Según Wright, debido a esto, emplear este modelo para explicar la semántica del lenguaje moral, todavía deja abierta la posibilidad de que sigamos sosteniendo el Realismo Moral, el cual es bastante implausible, como una *doctrina metafísica* (Cf. Wright 1992, pp. 8-9). Debo confesar que esta crítica me resulta completamente desconcertante. Al parecer, Wright cree que el Antirrealismo de Dummett es un modelo que permite explicar la semántica de un discurso; pero que no tiene implicaciones de carácter metafísico. En vista de esto, él llega a pensar que no hay inconsistencia alguna en ser Antirrealista Moral a la Dummett, en semántica, y a la vez ser Realista Moral, en metafísica. Sin embargo, creo que esta interpretación desconoce todas las afirmaciones de Dummett acerca de las implicaciones metafísicas que tiene el rechazo o la aceptación del Principio de Bivalencia en un discurso. En general, creo que está crítica se basa en el rechazo de la tesis fundamental de Dummett, de que las discusiones metafísicas se deben resolver en el terreno de la teoría del significado. De modo que, para que sus afirmaciones puedan ser tomadas en serio, Wright debe justificar su rechazo de esta tesis; y esto es algo que, por lo pronto, no he podido encontrar en su obra. No obstante, los planteamientos de Wright son importantes; en tanto que muestran que hay autores que se oponen a la idea de emplear el modelo Antirrealista dummettiano para explicar la semántica del lenguaje moral. En vista de ello, es necesario que presentemos al menos un argumento que nos permita justificar esta maniobra.

Creo que, para desarrollar tal argumento, podemos valernos de la misma estrategia que emplea Dummett para mostrar que el discurso acerca de colores es Antirrealista. Él parte de algunas consideraciones acerca de cómo el Realismo y el Antirrealismo tratan el fenómeno de la *vaguedad*. El Realismo, dada su aceptación del Principio de Bivalencia, debe comprometerse con la tesis de que todos los enunciados en los que usamos términos o expresiones vagas tienen un valor de verdad determinado. Así, desde esta perspectiva, los casos en los que no podemos determinar con certeza si un enunciado que contiene un término vago es verdadero o falso, deben interpretarse como casos en los que simplemente *no tenemos conocimiento* de los hechos relevantes que permiten saber cuál es el valor de verdad del enunciado. Naturalmente, esta incertidumbre puede superarse, si llegamos a conocer una serie de enunciados que nos den información suficiente acerca de esos hechos relevantes. De modo que, para el Realismo, el

fenómeno de la vaguedad debe tratarse como algo propio de nuestro lenguaje, de nuestras formas de expresión. En contraste, dado su rechazo del Principio de Bivalencia, el Antirrealismo puede explicar los casos en los que no podemos determinar con certeza el valor de verdad de un enunciado que contiene un término vago, diciendo simplemente que *la realidad es en sí misma vaga*; y que, en consecuencia, el valor de verdad de tales enunciados es esencialmente incierto. De este modo, para el Antirrealismo, el fenómeno de la vaguedad es algo que pertenece propiamente a la estructura del mundo (Cf. 1993a, p. 467). En vista de estas consideraciones, el primer paso en la estrategia de Dummett consiste en observar que los términos de color son vagos. Efectivamente, no es posible determinar con claridad la extensión de los predicados de color y el valor de verdad de muchos de los enunciados que los contienen es incierto. El segundo paso consiste en preguntarse qué implicaciones tiene, con respecto al discurso acerca de colores, adoptar un modelo Realista.

La respuesta de Dummett, a la cuestión de si podemos explicar de manera Realista la semántica del discurso acerca de colores, es que ello lleva inevitablemente a una paradoja. En efecto, desde el punto de vista del Realismo, el significado de un término como «rojo» debe especificarse diciendo que él se *refiere* a una propiedad real y determinada, a saber, la propiedad de ser-rojo. Así, cuando, debido a la vaguedad de este término, no podemos determinar cuál es el valor de verdad de un enunciado como «x es rojo»; lo que realmente ocurre es que desconocemos una serie de hechos relativos a la propiedad de ser-rojo, cuyo conocimiento nos permitiría determinar cuál es el valor de verdad del enunciado «x es rojo». Ahora, imaginemos una franja coloreada, que representa una parte del espectro de color. La mitad izquierda de la franja está coloreada de naranja y la mitad derecha está coloreada de rojo; sin embargo, la transición de un color a otro es tan gradual y continua, que entre un extremo y otro de la franja es posible apreciar una cantidad incontable de matices de naranja y rojo. Imaginemos, también, una serie de puntos numerados, dispuestos a igual distancia a lo largo de la franja. Estos puntos están tan cerca unos de los otros, que para el ojo humano el color sobre el que se encuentra cada punto resulta indistinguible del color sobre el que se encuentran los puntos aledaños. Según el Realista, dado que la propiedad de ser-rojo es determinada, debe ser posible ubicar un punto sobre la franja, tal que marquemos el límite entre la superficie de la franja que posee ésta propiedad y la superficie que no. Supongamos que sobre la franja hay 50 puntos y que el que marca este límite corresponde al marcado con el número 25. Pero, entonces, el Realista debe enfrentar un resultado paradójico. Por una parte, debe decir que la superficie que se encuentra entre los puntos 1 a 25 no posee la propiedad de ser-rojo; mientras la superficie que se encuentra entre los puntos 26 a 50 sí la posee. Pero por otra parte, debe reconocer que el color de la superficie en la que se encuentra el punto 25 es indistinguible del color de la superficie en la que se encuentra el punto 26; de modo que parece erróneo no atribuirle a la primera la propiedad de ser-rojo y sí atribuírsela a la segunda. Por consiguiente, en vista que lleva a una paradoja, el supuesto Realista de que el significado de «rojo» se especifica por referencia a una propiedad real y determinada debe ser rechazado. En general, en la medida en que para cualquier concepto vago podemos imaginar casos que lleven a este tipo de resultados paradójicos (de hecho, Dummett imagina casos que involucran conceptos matemáticos como «número grande»), debemos rechazar la idea de que la semántica de un discurso que contiene términos vagos puede especificarse de manera Realista. En contraste, frente a estos casos, el Antirrealismo simplemente dice que la realidad es en sí misma vaga; por lo cual, hay numerosos casos en los que no podemos decir con seguridad de entre los enunciados «x es naranja» y «x es rojo» cuál es verdadero y cuál es falso. Precisamente, ésta es la razón por la que parece más acertado especificar la semántica del discurso acerca de colores de manera Antirrealista (Cf. Dummett 1978b).

Como hemos dicho, podemos emplear esta misma estrategia para el caso del lenguaje moral. Lo primero que debemos observar es que los términos morales, efectivamente, son vagos. Al respecto, Russ Shafer-Landau ha afirmado:

[...] many moral concepts are "fuzzy around the edges," lack precise definitional criteria and so incompletely fix their extensions. In some cases, the puzzlement we experience about whether some act or quality of character truly exemplifies a moral property is due to this sort of vagueness, rather than some remediable cognitive limitation. Judgments ascribing the vague predicate to some action or state of affairs may be neither true nor false, owing to the extensional imprecision of the concept itself. (1994, pp. 333-4)

Así, si la semántica del lenguaje moral es Realista, entonces el significado de un término moral como «bueno» debe especificarse diciendo que él se *refiere* a una propiedad real y determinada, a saber, la propiedad de ser-bueno. De acuerdo con esto, la vaguedad del término «bueno» debe ser tratada de tal modo que, todos los casos en los que no podemos determinar cuál es el valor de verdad de un enunciado como «x es bueno» se conciben como casos en los que desconocemos una serie de hechos relativos a la propiedad de ser-bueno, los cuales son relevantes para determinar el valor de verdad del enunciado «x es bueno». Teniendo en cuenta esto, podemos imaginar una situación análoga a la que describe Dummett para el caso de los colores; pero que involucra el uso de conceptos morales. Observemos, por ejemplo, la siguiente secuencia de imágenes:



Cada imagen representa una acción que puede ser descrita como «abortar», practicada en etapas diferentes de la gestación humana. Es importante notar que, en este caso, la palabra «abortar» no está siendo usada en sentido propio. Estrictamente hablando, no es correcto emplear tal denominación, para referirse a acciones en las se emplean métodos anticonceptivos de emergencia (p.e., la píldora del día después) que impiden el desarrollo del cigoto. De tal forma, en este caso empleamos el término «abortar» de la manera más neutral posible, para referirnos a una acción que puede ser descrita de un modo más preciso como «eliminar *algo* que, como resultado de una unión sexual, ha comenzado a desarrollarse en el útero de una mujer». Ahora, podemos imaginar una secuencia mucho más amplia de imágenes, digamos 500, tales que la secuencia completa cubra todas las etapas de la gestación, desde el cigoto hasta un feto de aproximadamente 8 meses y medio. Adicionalmente, podemos suponer que la secuencia es tan completa y detallada que, entre una imagen y las que están más próximas a ella, no se aprecia ninguna diferencia significativa en las características de aquello (cigoto, embrión o feto) sobre lo que se practica la acción que se denomina «abortar». Finalmente, podemos suponer también que no hay ninguna discusión respecto al valor moral de algunas de las acciones representadas en la secuencia. Podemos decir sin problemas que, cuando el objeto del aborto es un cigoto, la acción es permisible; mientras que, cuando el objeto del aborto es un embrión o un feto, la acción es mala. Dadas estas condiciones, el Realista debe enfrentar un resultado paradójico. Supongamos que sólo en las imágenes 1 a 50 el objeto de la acción es un cigoto; mientras que en las imágenes 51 a 500 el objeto de la acción es primero un embrión y luego un feto. Entonces, por una parte, el Realista debe decir que

las acciones representadas en las imágenes 1 a 50 poseen la propiedad de ser-admisible; mientras que las acciones representadas en las imágenes 51 a 500 poseen la propiedad de ser-mala. Pero por otra parte, él también debe reconocer que el valor moral de la acción representada en la imagen 50 es indistinguible del valor moral de la acción representada en la imagen 51; de modo que parece erróneo atribuirle a la primera la propiedad de ser-admisible pero no atribuírsela a la segunda. Por consiguiente, en vista que lleva a una paradoja, el supuesto del Realista, según el cual el significado de términos como «bueno», «admisible» y «malo» se especifica por referencia a una propiedad real y determinada, debe ser rechazado.

En contraste, una explicación Antirrealista del lenguaje moral puede lidiar perfectamente con la vaguedad de los términos morales, sin caer en este tipo de paradojas. De acuerdo con tal explicación, la razón por la cual hay muchos casos, en los que no podemos decir con seguridad de entre los enunciados «abortar es admisible» y «abortar es malo» cuál es verdadero y cuál es falso, es que la *realidad moral*, en virtud de la cual determinamos el valor de verdad de los JMs, es en sí misma vaga. De tal modo, tal como ocurre en el caso de los colores, las consideraciones acerca de la semántica del lenguaje moral tienen implicaciones importantes respecto a la concepción de la metafísica que subyace al mismo. De acuerdo con esta posición, la realidad moral debe concebirse como algo que cobra existencia, gracias a la interacción entre rasgos objetivos y rasgos subjetivos propios de la condición humana; y no como algo que existe de manera determinada y con total independencia de la experiencia que los seres humanos podamos tener de ella. Como hemos visto, esta dependencia de la experiencia es precisamente la que le da ese carácter indeterminado, que hace que los términos morales sean esencialmente vagos. Así, tenemos al menos un argumento que nos permite sostener la idea de que el discurso moral es Antirrealista en el sentido de Dummett; lo cual, como hemos visto, brinda una alternativa frente a las posiciones Cognitivistas tradicionales, como el Intuicionismo y el Realismo Moral, y frente a las posiciones Antirrealistas subjetivistas, como el Expresivismo y las Teorías del Error.

Para terminar, es importante darnos cuenta de que sostener esta posición inevitablemente nos obliga a considerar un par de cuestiones importante. La primera de ellas tiene que ver con la idea de que el Antirrealismo de Dummett requiere de la noción de un método objetivo de justificación. Así, en el caso del lenguaje moral, ¿cuál sería ese método? Las consideraciones que hemos hecho hasta acá muestra una analogía importante entre el discurso acerca colores y el lenguaje moral. Por tal razón, podemos responder a este interrogante, diciendo que tal método corresponde a una *experiencia moral*. Sin embargo, en el capítulo 1, ya vimos los inconvenientes que presenta es idea. Por tal razón, si queremos darle un poco de plausibilidad a la posición que ahora estamos asumiendo, debemos darnos a la tarea de especificar claramente en qué consiste la experiencia moral. De otro lado, la segunda cuestión que debemos considerar tiene que ver con algo en lo que hemos venido insistiendo constantemente, a saber, que una teoría que pretenda explicar satisfactoriamente el lenguaje moral necesariamente debe dar cuenta del Internalismo. Entonces, suponiendo que podemos sostener sin problemas la idea de que la experiencia moral constituye un método objetivo de justificación, ¿esto cómo da cuenta del Internalismo? Creo que a esta cuestión también se puede responder, a partir de una especificación clara de en qué consiste la experiencia moral. El siguiente numeral estará dedicado a cumplir con esta tarea.

### **3.4. La Experiencia de Intersubjetividad**

Como dijimos en la sección 1.4, la Psicología Humeana no debe ser concebida simplemente como una doctrina psicológica. Ella está ligada inevitablemente a una tesis

metafísica, a saber, la tesis de que el mundo es motivacionalmente inerte y está constituido por propiedades del estilo de las cualidades primarias de Locke. Las afirmaciones que hemos hecho hasta acá, dejan ver con claridad que esta doctrina constituye el trasfondo de la mayor parte de las teorías metaéticas. Por la misma razón, podemos decir que ella también está a la base de las críticas que tradicionalmente se han hecho a la concepción Intuicionista de la experiencia moral. En efecto, si admitimos tal doctrina, nuestra percepción debe ser vista como una facultad que permite la captación de rasgos del mundo, los cuales existen con total independencia de nuestra mente. De tal modo, desde esta perspectiva, sólo podemos interpretar a la experiencia moral de dos formas. O bien, podemos verla como una experiencia *puramente subjetiva* que carece por completo de valor cognitivo, esto es, como la manifestación de ciertas emociones o sentimientos propios del sujeto. O bien, podemos verla como el producto del ejercicio de una *facultad especial*, en virtud de la cual captamos ciertas propiedades especiales que hay en el mundo. Tal como lo veo, estas dos interpretaciones son erróneas, principalmente por que se basan en una doctrina errónea: en la Psicología Humeana. Demostrar efectivamente que ella es errónea es una tarea mayúscula; una tarea que, para ser llevada a cabo, requiere, entre otras cosas, abandonar el terreno de la metaética y adentrarse en los dominios propios de la teoría de la acción. Ciertamente esto excede los propósitos de este trabajo, y por tal razón, no me ocuparé de esta cuestión aquí. Simplemente digamos que los argumentos propuestos por McDowell en contra de las posiciones No-Cognitivist, los cuales reseñamos en el capítulo 2, nos dan buenas razones para, por lo menos, considerar sospechosa a la Psicología Humeana. En vista de ello, daré por sentado que ella es errónea. Una consecuencia que se sigue de esto, es que probablemente hay otras formas de entender a la experiencia moral diferentes a las dos que hemos indicado. En vista de ello, el propósito de este numeral es esbozar una concepción de la experiencia moral, que no haga de ella algo puramente subjetivo ni la comprometa con extravagancias epistemológicas y metafísicas. Esta concepción corresponde a lo que Gallagher ha denominado *Intersubjetividad Primaria* y, como veremos, ella está respaldada por un número importante de investigaciones de carácter empírico, adelantadas principalmente en el campo de la psicología del desarrollo.

La interpretación de la experiencia moral que quiero proponer se basa en la idea de que, desde el momento de nacer, tenemos la capacidad de reconocernos a nosotros mismos como sujetos, de reconocer a otras entidades en nuestro entorno como sujetos y de reconocer a esas entidades como similares a nosotros. Gracias al ejercicio de esta capacidad, tenemos conocimiento de los otros, sin necesidad de hacer simulaciones ni de hacer inferencias a partir de una teoría de la mente. Más bien, tenemos tal conocimiento de manera directa; en virtud de que, en cierto sentido, experimentamos las emociones e intenciones de los otros. Como hemos dicho, Gallagher denomina a esta experiencia Intersubjetividad Primaria (2001, p. 86). Ella es concebida como primaria, no sólo porque aparece en una etapa temprana del desarrollo; sino, principalmente, porque a lo largo de nuestras vidas ella es la herramienta primordial de que disponemos, para entender a los otros y para involucrarnos en prácticas que requieren cooperación (Gallagher & Zahavi 2008, p. 189). Adicionalmente, esta experiencia se sustenta en la percepción multimodal, en la capacidad que tenemos de integración de información proveniente de las diferentes modalidades sensoriales de nuestro aparato perceptual. Gracias a ello, es posible atribuirle esta experiencia a alguien, sin necesidad de comprometerse con la idea de que esa persona posee alguna facultad especial. Pero, ¿exactamente en qué consiste esta experiencia? y ¿qué evidencia hay de que efectivamente la poseemos desde una edad muy temprana? Para responder a estas cuestiones es necesario considerar algunas investigaciones de carácter empírico.

La evidencia empírica es bastante numerosa, así que necesariamente mencionaremos sólo unos cuantos estudios. En primer lugar, varios estudios sugieren que los bebés de 5 meses de nacidos tienen la capacidad de reconocer emociones en personas adultas; pero este reconocimiento requiere de la percepción multimodal. Así, por ejemplo, A.S. Walker-Andrews y Lennon (1991) reportan que su investigación mostró que la capacidad de los bebés de discriminar expresiones vocales que indicaban felicidad o tristeza, requiere de la presencia de las expresiones faciales que usualmente las acompañaran. En su estudio, se pudo observar que, cuando a un bebé se le presenta el sonido de una voz que evidenciaba alguna emoción, acompañado de una imagen que representaba la expresión facial correspondiente a tal emoción; él respondía adecuadamente a la misma. Lo cual indica que él es capaz de distinguirla. En contraste, los bebés a los que sólo se les presentó un estímulo auditivo no dieron muestra de reconocer la emoción que se expresaba mediante él.

Adicionalmente, otros estudios han mostrado que los bebés tienen la capacidad de imitar las expresiones verbales y faciales de adultos. Kessen, Levine y Wendrich (1979), por ejemplo, mostraron que, desde la edad de 3 meses, los niños tienen la capacidad de imitar el tono de los sonidos vocales proferidos por adultos. Estos investigadores hicieron grabaciones de interacciones verbales, en las que se escuchaba la voz del adulto con determinada tonalidad, seguida de la voz del niño. Las grabaciones fueron analizadas por expertos en audio, quienes determinaron que la correspondencia en las tonalidades de las dos vocalizaciones superaba el 50%. Un experimento similar fue llevado a cabo por Legerstee (1990). En tal experimento se formaron dos grupos de bebés de 4 meses de edad. Al primer grupo se le presentaron grabaciones en las que se pronunciaba una vocal, acompañadas de la imagen de una persona pronunciando la misma vocal que en el audio; mientras que al segundo grupo se le presentaron las mismas grabaciones, pero acompañadas de la imagen de una persona pronunciando una vocal diferente a la del audio. Posteriormente, se grabaron las preferencias que hacían los bebés, tras escuchar las grabaciones. Se observó que los niños del grupo al que se le presentaron audios e imágenes concordantes lograron imitar el sonido; mientras que los niños del otro grupo casi nunca pudieron imitarlo. Así, este último estudio muestra que la capacidad de imitar sonidos depende completamente de la percepción multimodal.

Por otra parte, algunos estudios muestran que la imitación no se reduce a preferencias vocales. Vitner (1986), siguiendo algunos experimentos hechos originalmente por Meltzoff y Moore (1983), mostró que los recién nacidos pueden imitar acciones, como abrir la boca o sacar la lengua. Su estudio encontró una particularidad: los bebés pueden imitar tales acciones, cuando se les presentan imágenes dinámicas, p.e., un video de un adulto abriendo la boca; pero no, cuando se les presentan imágenes estáticas, como la foto de un adulto con la boca abierta. Legerstee (1991) señala un resultado aún más particular. En su estudio, se pudo observar que los niños de 5 a 8 meses sólo imitan acciones, cuando éstas son realizadas por otros seres humanos. En las situaciones en las que se les presentaba a un objeto inanimado realizando una acción, los niños no imitaban esta acción. A partir de estos resultados, Eleanor Gibson y Anne Pick han señalado que la imitación sólo ocurre en una situación comunicativa, propia de un *contexto social* (Cf. 2000, p. 62).

Estas consideraciones de Gibson y Pick pueden apoyarse a partir de una serie de estudios en los que se examina lo que suele denominarse interacción cara a cara. En sus experimentos, Cohn y Tronick (1987, 1988) observaron interacciones cara a cara entre bebés de 3 meses de edad y sus madres. A partir de las observaciones, se pudo establecer que, en los intercambios comunicativos que involucraban expresiones de emociones, los niños asumen expresiones positivas, luego de percibir las en sus madres. En un segundo experimento, se observaron interacciones cara a cara que involucraban niños de entre 6 y

9 meses de edad. Se pudo determinar que, en tales situaciones, los niños estaban dispuestos a asumir expresiones positivas, incluso antes que sus madres lo hicieran. Estos resultados sugieren que, en la interacción cara a cara, las acciones de los niños siempre están reguladas afectivamente. Es decir, que ellas se producen dependiendo de las expresiones emotivas que ellos perciben en su entorno.

Estas investigaciones evidencian que la Intersubjetividad Primaria es un tipo de experiencia que nos vincula directamente con otros seres humanos. Como hemos dicho, gracias a ella nos reconocemos como sujetos y reconocemos a otras entidades de nuestro entorno como sujetos similares a nosotros. Es importante notar que, cuando tenemos esta experiencia, reconocemos a los otros, gracias a que podemos experimentar lo que ellos están experimentando. Esto se puede apreciar en los experimentos que involucran imitación e interacción cara a cara. En efecto, en tales situaciones, parece como si los niños experimentaran como propios los movimientos y emociones de los adultos. Adicionalmente, estas afirmaciones pueden encontrar sustento en estudios empíricos provenientes de otra disciplina. Varios estudios neurocientíficos han mostrado que el cerebro humano posee lo que puede denominarse un *sistema de neuronas en espejo*. Este sistema involucra principalmente áreas de la corteza pre-motora. Allí, las neuronas se activan, tanto en las situaciones en las que el sujeto está realizando una acción de una clase determinada, como en las situaciones en las que el sujeto percibe que otra persona realiza una acción de la misma clase. Así, por ejemplo, Rizzolatti y Arbib (1998) han notado que el área de Broca, la cual se asocia tradicionalmente al habla, también está asociada a la ejecución de movimientos con la mano y el brazo. Gracias a tomografías de emisión de positrones (conocidas como PET), Rizzolatti y Arbib han observado que parte del área de Broca pertenece al sistema de neuronas en espejo; pues, allí las neuronas se activan, tanto en los casos en que el sujeto ejecuta movimientos con la mano y el brazo, como en los casos en que el sujeto percibe que otra persona ejecuta la misma clase de movimientos. A partir de esto, ellos han concluido que el sistema de neuronas espejo hace posible una forma de comunicación más primitiva que el habla, puesto que el área de Broca, la cual es esencial para la comunicación verbal, se ha desarrollado sobre una parte de ese sistema. Por otra parte, Gallese, Keyses y Rizzolatti (2004) han mostrado que el sistema de neuronas espejo también abarca áreas vinculadas a la experiencia emotiva. Ellos han observado que pacientes con daños en la ínsula, una región del subcortex asociada al reconocimiento de expresiones faciales de asco, pierden la capacidad de experimentar asco. En efecto, tales pacientes pueden ingerir alimentos incomedibles de manera indiscriminada, así como ver fotografías que representan comida repugnante, sin experimentar ninguna sensación de asco. Así, la evidencia sobre la existencia de un sistema de neuronas espejo, sugiere que, en efecto, la Intersubjetividad Primaria nos permite experimentar de manera directa las emociones y las intenciones de los otros.

Por otra parte, Gallagher y Zahavi han afirmado que la Intersubjetividad Primaria es una experiencia intermodal que tenemos desde el nacimiento. Ella nos permite conocer las posibilidades de acción que nos ofrece nuestro propio cuerpo, así como las oportunidades de acción que otros seres humanos nos ofrecen (Cf. 2008, p. 188). De acuerdo con esto, un aspecto importante de esta experiencia es que ella posee un *vínculo con la acción*. Como Gallagher y Zahavi han señalado, esto se debe precisamente al hecho de que la percepción es un fenómeno enactivo, que involucra procesos motores. Así, al valerse de las diferentes modalidades de nuestra percepción sensorial, la experiencia intersubjetiva también adquiere ese carácter enactivo. De ahí que se diga que ella nos presenta oportunidades de acción, las cuales están determinadas, tanto por nuestro propio cuerpo, como por aspectos corporales y emocionales de los otros (Cf. *Idem*, p. 179). Esto puede expresarse de otro modo, a la luz de algunas afirmaciones que



J.J. Gibson ha hecho, en su famoso libro *The Ecological Approach to Visual Perception*. Allí, él afirma:

[The] bodies are subjects to the laws of mechanics and yet *not* subjects to the laws of mechanics, for they are not *governed* by these laws. They are so different from ordinary objects that infants learn almost immediately to distinguish them from plants and nonliving things. When touched they touch back, when struck they strike back, in short, they *interact* with the observer and with one another. Behavior affords behavior [...]. Sexual behavior, nurturing behavior, fighting behavior, cooperative behavior, economic behavior, political behavior –all depend on the perceiving of what another person or other persons afford, or sometimes on misperceiving of it. [...] The perceiving of these mutual affordances is enormously complex, but it is nonetheless lawful, and it is based on the pickup of the information in touch, sound, odor, taste, and ambient light. (1986, p.135)

Así, de acuerdo con esta lectura, podemos entender a la Intersubjetividad Primaria como un tipo de experiencia, en la que al sujeto se le presenta unas *affordances* que poseen significación social (Cf. Gibson y Pick 2000, p. 53-74).

Este es tal vez el aspecto que más me interesa de la Intersubjetividad Primaria, puesto, de acuerdo con él, esta experiencia es esencialmente motivacional. De este modo, si concebimos a la experiencia de intersubjetividad como el método de justificación objetivo que una explicación Antirrealista del lenguaje moral requiere; entonces las condiciones de justificación de los JMs están necesariamente ligadas a una normatividad externa al sujeto, la cual regula sus acciones en un contexto social. La objetividad de esta experiencia viene dada por el hecho de que ella se constituye, a partir de rasgos inherente a la condición humana. Adicionalmente, dado su carácter enactivo, esencialmente motivacional; una teoría Antirrealista que recurra a ella, para explicar la semántica del lenguaje moral, podrá dar cuenta sin problema alguno del Internalismo. Creo que esta teoría aún debe desarrollarse de mucho más amplia; sin embargo, creo que el esbozo que en este capítulo hemos hecho de ella es suficiente para darnos cuenta de su plausibilidad y para marcar el camino para una investigación futura.

## 4. La Polémica Internalismo vs Externalismo

Hasta ahora, hemos dado por supuesto que el Internalismo, entendido como la tesis general de que existe una conexión interna y necesaria entre los JM y nuestras motivaciones para actuar, es verdadero. A partir de este supuesto, hemos sostenido que una teoría metaética que pretenda explicar satisfactoriamente la semántica del lenguaje moral necesariamente debe dar cuenta de él. Precisamente, ésta fue una de las razones por las que consideramos, en el capítulo 3, que la teoría expuesta allí es la única que resuelve de manera satisfactoria los problemas que han inquietado a los teóricos en la metaética contemporánea. En vista de ello, en este capítulo nos introduciremos en la discusión acerca del Internalismo. Como hemos dicho, la polémica entre Internalistas y Externalistas ha sido una de las más intensas en la metaética contemporánea. Por tal razón, si queremos convencernos de que el Internalismo es verdadero, las observaciones de sentido común acerca del carácter motivacional de los JMs no parecen suficientes. Debemos, entonces, realizar una defensa del Internalismo, frente a las objeciones de que ha sido objeto por parte de los Externalistas. El propósito de este capítulo es, justamente, presentar esta defensa.

### 4.1. Dos Versiones del Internalismo

Hemos caracterizado al Internalismo, de manera general, como la doctrina según la cual existe una conexión *interna* y *necesaria* entre los JMs y nuestras motivaciones para actuar. Según esto, para los Internalistas, los JMs son intrínsecamente motivacionales; o en otras palabras, ellos constituyen motivos o razones para actuar sólo en virtud de su *contenido*. Así, una persona que acepte sinceramente el Juicio «x es buena», en donde x es una acción cualquiera, y tenga plena comprensión del significado de tal Juicio, estará motivado a hacer x, sólo en virtud de su aceptación del mismo. En contraste, el Externalismo es la doctrina según la cual la conexión entre los JMs y nuestras motivaciones es *externa* y *contingente*; de modo que, los JM son incapaces por sí mismos de motivarnos a realizar acciones. De acuerdo con los Externalistas, el que un JM nos motive a actuar es algo que depende, tanto del contenido de ese Juicio, como de *hechos adicionales* acerca los agentes, de sus preferencias y de sus razones para actuar. Ahora, si bien la discusión puede plantearse a partir de esta caracterización general de ambas posiciones, es importante señalar que, entre los propios Internalistas, hay una discusión acerca de cómo debe entenderse la afirmación de que existe una conexión interna entre JMs y motivaciones. En este numeral examinaré dos interpretaciones distintas de esta afirmación, que pueden encontrarse en la literatura contemporánea.

La primera aclaración que debe hacerse, a fin de apreciar las distintas formas en que puede ser sostenido el Internalismo, parte de la siguiente consideración. A primera vista, el Internalismo parece una doctrina bastante implausible; ya que pueden pensarse muchas situaciones en las que un agente juzga sinceramente que «x es buena», tiene plena comprensión del Juicio y, no obstante, ciertas *consideraciones prudenciales*, esto es., consideraciones acerca de su interés personal, lo inclinan a no hacer x. Así, por

ejemplo, puede imaginarse una situación en la que un agente juzga que es bueno dar dinero para ayudar a los pobres y, no obstante, debido a que en la actualidad sus recursos económicos son bastante limitados, tal agente carece por completo de la motivación de dar dinero. Sin embargo, estas consideraciones no representan un problema mayor para el Internalismo. Como hemos dicho, la afirmación fundamental de éste es que los JM son intrínsecamente motivacionales, que motivan sólo en virtud de su contenido; pero esto no implica que él deba comprometerse con la afirmación mucho más fuerte de que, siempre que un agente sostiene un JM, en virtud de ello, posee un motivo para actuar, cuya fuerza es superior a la de cualquier otro motivo no-moral. En otras palabras, asumir el Internalismo no implica negar que los JMs están sometidos a procesos deliberativos y que, en determinadas circunstancias, el resultado de esos procesos deliberativos puede llevar a que otras motivaciones se impongan a la motivación moral y sean mucho más efectivas respecto a la acción que ésta. Michael Smith expresa esto diciendo que el Internalismo debe concebir a los JMs como razones *pro tanto*, esto es, como estados psicológicos que tienen poder motivacional sólo cuando, como resultado de un proceso de deliberación racional, ellos se presentan como las razones más fuertes para actuar. Así, para no caer presa de la objeción que hemos venido considerando, cualquier formulación del Internalismo debe valerse de algún recurso que le permita dar cuenta de ese carácter especial de la motivación inherente a los JMs. Según Smith, este recurso es la introducción de una cláusula *ceteris paribus* en cualquier formulación del Internalismo (Cf. Smith, 1994, pp. 6-7).

A la luz de estas consideraciones, ahora podemos decir que la primera interpretación del Internalismo considera que lo que hay detrás de él puede expresarse del siguiente modo:

**(Int1):** Si un agente juzga que para él es correcto  $\phi$ ar en las circunstancias C, entonces *ceteris paribus* él está motivado a  $\phi$ ar en C

Esta versión del Internalismo suele asociarse a autores como John McDowell (1998) y Mark Platts (1979). Sin embargo, varios autores han cuestionado esta interpretación del Internalismo. Según ellos, es posible imaginar situaciones en las que el proceso deliberativo del agente ha privilegiado al JM sobre otras consideraciones de tipo prudencial, es decir, en las que se satisface la cláusula *ceteris paribus*; y en las que, no obstante, el agente carece por completo de la motivación para actuar de conformidad con su JM. La conducta de personas con algún tipo de adicción o con una severa depresión ejemplifican este tipo de situaciones. Un adicto a las drogas, por ejemplo, puede juzgar *ceteris paribus* que no debe drogarse, y no obstante, debido a la gravedad de su adicción, no tener ninguna motivación de dejar de drogarse. De igual modo, una madre con una severa depresión puede juzgar *ceteris paribus* que debe cuidar de sus hijos, y no obstante, debido a la intensidad de su depresión, no tener ninguna motivación de efectivamente cuidar de ellos. Quienes se comportan de este modo exhiben debilidad de voluntad. De modo que, el problema para el **Int1** es que él está formulado de un modo tal, que no permite que haya casos de debilidad de voluntad. Por ello, en la medida en que en la vida cotidiana encontramos múltiples ejemplos de debilidad de voluntad, el **Int1** resulta insostenible (Cf. Smith, 1994, p. 61; Wedgwood, 2007, p. 26).

En vista de que la debilidad de la voluntad es una forma de irracionalidad práctica, Smith y Wedgwood han propuesto una versión un poco más débil del Internalismo, en la que se incluye una cláusula relativa a la racionalidad del agente. Así, según ellos, lo que hay detrás del Internalismo puede expresarse del siguiente modo (Smith *Idem*; Wedgwood *Idem*):

**(Int2):** Si un agente juzga que para él es correcto  $\phi$ ar en las circunstancias C y él es racional, entonces *ceteris paribus* él está motivado a  $\phi$ ar en C.

Esta segunda interpretación del Internalismo parece mucho más plausible, ya que no está expuesta a contraejemplos basados en la debilidad de la voluntad, como sí lo está la primera interpretación. Sin embargo, considero que esta Interpretación es mucho más problemática que la primera y no da cuenta de lo que propiamente constituye al Internalismo. En lo que queda de esta sección trataré de justificar esta afirmación.

Un primer inconveniente con **Int2** tiene que ver con la noción de racionalidad que se emplea en él. Adina Roskies ha argumentado que, a menos que los defensores de esta interpretación expliquen satisfactoriamente de en qué consiste ser prácticamente racional, esta versión del Internalismo será trivialmente verdadera. En efecto, hasta ahora, los defensores de **Int2** han tomado como ejemplos paradigmáticos de irracionalidad práctica los casos de personas que, como los adictos y las personas con depresión severa, no actúan de acuerdo con lo que juzgan. Según esto, para ellos, ser racional consiste en *actuar de acuerdo con lo que se juzga*. Pero, si esta es la noción de racionalidad que está presente en **Int2**, entonces cualquier caso que podamos imaginar, en el que un agente no actúa de acuerdo con sus Juicios, puede tratarse como un caso en el que no simplemente no se cumple la cláusula de racionalidad, de un modo tal, que nunca pueda cuestionarse la verdad de **Int2**. Esto es, justamente, lo que hace a esta interpretación trivialmente verdadera (Cf. Roskies 2003, p. 53). Para responder a esta crítica, Smith y Wedgwood han tratado de especificar en qué consiste ser prácticamente racional, a partir de la noción de “requisitos diacrónicos de racionalidad”. Según ellos, un agente es prácticamente racional, cuando tiene la disposición a hacer que sus acciones, pensamientos y motivaciones se ajusten a las reglas y procedimientos propios del razonamiento, esto es, cuando tiene la disposición a mantener cierta *coherencia interna* entre sus acciones, pensamientos y motivaciones. Según ellos, esta noción de racionalidad salva al **Int2** de ser trivialmente verdadero; puesto que le deja cierto margen de maniobra al Externalista; en efecto, éste puede tratar de imaginar una situación en la que un agente es racional (en el sentido que se acaba de indicar), juzga *ceteris paribus* que cierta acción es buena y, no obstante, no actúa de conformidad con su Juicio. El defensor de **Int2**, sin embargo, insistirá en que tal situación realmente no es posible (Smith 2008, pp. 213-5; Wedgwood 2007, pp. 26-7).

Pero estas aclaraciones acerca de la noción de racionalidad que se emplea en la formulación de **Int2** no despejan todas las dudas que hay en torno a esta interpretación. En efecto, Roskies ha señalado que **Int2** realmente no indica nada específico del Internalismo; ya que una posición Externalista puede perfectamente dar cuenta de los requisitos que en tal enunciado se presentan. Como hemos dicho, según el Externalismo, el que un JM tenga poder motivacional depende, no sólo de su contenido; sino también de hechos adicionales acerca del agente, de sus preferencias y de sus razones para actuar. Como veremos, de acuerdo con esta doctrina, un JM motiva, siempre y cuando el agente tenga un *deseo adicional* de realizar las acciones que evalúa como buenas y evitar las que evalúa como malas. Así, el Externalista puede admitir sin problema que, todo agente que posea este deseo adicional, sea racional y juzgue *ceteris paribus* que cierta acción es buena; necesariamente estará motivado a realizar esa acción. De igual modo, el Externalista puede admitir que, todo agente que posea este deseo adicional, juzgue *ceteris paribus* que cierta acción es buena y, no obstante, carezca de la motivación de realizar esa acción; necesariamente debe ser considerado como prácticamente irracional (Cf. Roskies 2003, p. 53). Seguramente el defensor de **Int2** insistirá en que, de todos modos, aún es posible distinguir al Internalismo (tal como él lo interpreta) del Externalismo. Esta diferencia puede trazarse del siguiente modo: mientras que, para el

Internalista, *todo* agente que no actúa de conformidad con sus JMs es prácticamente irracional; para el Externalista, sólo *algunos* de los agentes que no actúan de conformidad con sus JMs lo son, a saber, aquellos que poseen el deseo adicional de realizar las acciones que evalúa como buenas y evitar las que evalúa como malas. Sin embargo, creo que quien responde de este modo pierde de vista aquello a lo que Roskies apunta. Como lo veo, el punto de Roskies es, más bien, que las cuestiones acerca de la racionalidad son irrelevantes al considerar el carácter motivacional de los JMs. De hecho, al introducir la cláusula de racionalidad en **Int2**, los defensores de esta interpretación formulan un enunciado que parece caracterizar mejor al Externalismo que al Internalismo. En efecto, de acuerdo con **Int2**, la motivación moral es algo que depende principalmente de la racionalidad del agente. Ya que la racionalidad ha sido concebida como una disposición a mantener cierta coherencia interna; desde este punto de vista, la motivación moral no puede pensarse como algo inherente a los JMs, algo ligado a ellos en virtud de su contenido; sino como algo que depende de un hecho adicional acerca del agente, a saber, que él posea justamente esa disposición. Y, como hemos dicho, esto último es lo que precisamente caracteriza al Externalismo: el tener que recurrir a hechos adicionales acerca del agente para explicar la motivación moral (Cf. *Idem*, p. 54-5).

A partir de lo anterior, podemos concluir que **Int2** no es una buena interpretación del Internalismo. Como lo veo, esto implica que debemos quedarnos con **Int1**, como la mejor interpretación del Internalismo, y que debemos tratar de lidiar con los contraejemplos basados en la debilidad de voluntad. Creo que esto se puede hacer de dos formas. En primer lugar, Davidson (1970) ha sugerido que, lo que ocurre en las situaciones en que un agente parece exhibir debilidad de voluntad, es que ese agente realmente no conoce con claridad cuáles son sus motivos para actuar. Este desconocimiento lo lleva a atribuirse motivos que en verdad no posee o que, aún cuando sí los posee, no son sus motivos para actuar más fuertes. A partir de esta concepción de la debilidad de voluntad, podemos pensar que lo que realmente ocurre en casos como el del adicto, es que el agente ha intentado comprometerse con el Juicio de que no debe drogarse, sin llegar a darse cuenta de que la fuerza motivacional de su deseo de drogarse es mucho más fuerte que la fuerza motivacional de ese Juicio. En otras palabras, podemos concebir a este tipo de casos como situaciones en las que no se cumple la cláusula *ceteris paribus* y el agente no logra percatarse de ello. En segundo lugar, algunos autores piensan que el problema con **Int1** es que, de acuerdo con él, el compromiso de un agente con un JM determinado, lo lleva inevitablemente a actuar de conformidad con ese JM. Quienes piensan de este modo, conciben a los motivos como *mecanismos causales* que, de un modo más o menos determinista, llevan a la acción. Por tal razón, para estos autores, **Int1** es insostenible, ya que no deja cabida a la debilidad de voluntad. Sin embargo, esta concepción general de los motivos es errónea. Los motivos deben ser vistos, más bien, como estados psicológicos que nos *inclinan* a actuar de un modo determinado, cuya atribución a un agente permite *explicar* las acciones del mismo (Cf. Smith 1994, pp. 101-4). Esta concepción de los motivos deja abierta la posibilidad de que un motivo, que es el más fuerte en la psicología del agente, pueda tener tan poca fuerza, que es incapaz de llevar efectivamente a la acción; lo cual puede ocurrir debido a cuestiones externas a la psicología del agente, como por ejemplo, a una condición neurofisiológica específica. Para ponerlo en otros términos, lo único que el Internalista necesita, es sostener la tesis de que los JMs son intrínsecamente motivacionales; no hay nada que lo obligue a comprometerse con una tesis específica acerca de qué tanta fuerza motivacional tienen los JMs o qué tan efectivos son para causar las acciones de un agente. Una vez aclarado esto, podemos tratar los contraejemplos de personas con una depresión severa, como situaciones en las que el agente, realmente no carece de motivación; sino que posee un motivo cuya fuerza es demasiado débil, debido a cierta condición neurofisiológica. De este

modo, podemos sostener Int1, sin preocuparnos por los casos en que hay debilidad de voluntad: tales casos realmente no son contraejemplos para esta interpretación.

Así, a partir de lo anterior, podemos decir que hay buenas razones para considerar a **Int1** como la mejor formulación del Internalismo. En lo que sigue, daré por cerrada la discusión de que nos hemos ocupado en este numeral y emplearé la denominación general "Internalismo", para referirme específicamente a la versión **Int1**.

## 4.2. La defensa de Hare del Internalismo

En la discusión contemporánea entre Internalistas y Externalistas es posible distinguir dos momentos. El primero de ellos comprende el periodo que va desde los años 50s hasta los años 90s del siglo XX. En este periodo la discusión se da a un nivel puramente conceptual; para ponerlo en los términos de una fórmula que últimamente se usa con cierta frecuencia, en este periodo la discusión tiene el carácter propio de la "filosofía de sillón". El segundo momento corresponde a los años que han corrido del siglo XIX. En este periodo la discusión se ha caracterizado por recurrir a estudios de carácter empírico, ajenos a la filosofía, propios principalmente de las neurociencias. En este trabajo me ocuparé únicamente de la discusión correspondiente al primer periodo, la discusión netamente filosófica. La discusión correspondiente al segundo periodo, como hemos dicho, involucra una serie de consideraciones provenientes de las ciencias empíricas, las cuales, debo confesar, no logro comprender plenamente. Por tal razón, he optado por omitir cualquier consideración respecto a ella.

Como hemos dicho, la primera defensa importante del Internalismo fue hecha por R.M. Hare en su libro *The Language of Morals* (1952). En el capítulo 2 vimos que, para Hare, el significado del término «bueno» tienen dos componentes: uno descriptivo y el otro prescriptivo. El componente descriptivo tiene que ver con los estándares que convencionalmente empleamos para determinar qué acciones cuentan, en nuestro contexto histórico-social, como buenas. Mientras que el componente prescriptivo tiene que ver con el hecho de que, independientemente de qué estándares se empleen en un contexto determinado para determinar qué acciones son buenas, «bueno» siempre se emplea para *prescribir* o *recomendar* acciones. La idea de que el significado de «bueno» posee este elemento prescriptivo es, precisamente, lo que permite sostener el Internalismo; ya que, en virtud tal elemento, los JMs pueden constituir motivos o razones para actuar. No obstante, Hare reconoce que hay situaciones en las que parece que un agente admite, *ceteris paribus*, un JM determinado y, no obstante, carece por completo de la motivación de actuar conforme a tal Juicio. El que podamos imaginar situaciones de este tipo parece una razón de peso para negar su idea de que el significado de los términos morales está constituido por un elemento prescriptivo. Hare, sin embargo, piensa que lo que realmente ocurre en tales situaciones es que el agente está usando el término moral en lo que él denomina "comas invertidas":

There are, however, cases in which we use the Word "good" with no commendatory meaning at all. We must distinguish several kinds of such non-commendatory use. The first has been called the *inverted-commas* use. [...] We are, in this use, not making a value-judgement ourselves, but alluding to the value-judgement of other people. (1952, p. 124)

Según esto, es perfectamente posible que exista un agente, que juzgue sinceramente que para él « $\phi$ ar es bueno», sin que, por ello, tenga la motivación de  $\phi$ ar. Este comportamiento es propio de sociópatas o amoralistas, los cuales son capaces de reconocer que ciertas acciones, por ejemplo, torturar a un ser humano, son malas; aún cuando, ellos carecen por completo de la motivación para no torturar (o para dejar de torturar) a otro ser

humano. El punto de Hare es que, en casos como estos, no se puede decir propiamente que el agente juzga que  $\phi$ ar es bueno; sino que éste debe ser concebido como alguien que juzga que  $\phi$ ar es una de las cosas que *otras personas, pero no él*, consideran bueno. Así, el Juicio del sociópata de que «torturar es malo», de acuerdo con Hare, debe entenderse propiamente como el Juicio de que «para otras personas (pero no para mí) torturar es malo». Esta es la razón por la que estos sujetos carecen por completo de la motivación de actuar de conformidad con sus JM. En efecto, en la medida en que entienden el término «bueno» en comas invertidas, lo despojan del elemento prescriptivo, propio de su significado. En la medida en que, como hemos dicho, este elemento es constitutivo del significado de los términos morales; debemos decir que, en sentido estricto, los sociópatas o amoralistas no emplean conceptos morales en sus Juicios. Y si estos sujetos se comportan sistemáticamente de este modo, entonces debemos concebirlos como sujetos que en realidad nunca usan conceptos morales. Según Hare, a partir de esto podemos considerar a un sociópata como una persona que *no comprende plenamente* el significado de los conceptos morales; ya que su comprensión de términos como «bueno» no involucra el elemento prescriptivo que es esencial a la semántica de los mismos. Si bien el sociópata tiene pensamientos y profiere enunciados que tienen la forma de JMs, en realidad él tiene pensamientos y profiere enunciados en los que los conceptos morales se usan en comas-invertidas; por ello, él carece por completo de la motivación para actuar de conformidad con tales Juicios (Cf. Hare 1952, pp. 111-26).

A partir de estas consideraciones, podemos darnos cuenta de que el Internalismo efectivamente es verdadero. Debido a que, siempre que un agente tiene plena comprensión de los términos morales, es imposible que él juzgue que cierta acción es buena y, a la vez, carezca de la motivación para actuar de conformidad con su Juicio; debemos decir que necesariamente la motivación es algo intrínseco a los JMs mismos, algo que está ligado a ellos únicamente en virtud de su contenido. Esto último es, precisamente, lo que afirma el Internalismo.

### **4.3. El Amoralismo de D.O. Brink y el Externalismo**

David O. Brink fue el primero en atacar el argumento de Hare, sobre el uso de los términos morales en comas-invertidas, y en tratar de darle plausibilidad al Externalismo. El punto de partida de Brink es el Realismo Moral, esto es, la idea de que los JMs son creencias, que poseen valor de verdad y describen estados de cosas en el mundo. Brink es uno de esos autores que considera que el Realismo Moral es clave para dar cuenta de las pretensiones de objetividad de la ética. Como hemos visto, la idea de objetividad involucra la creencia de que hay JMs correctos e incorrectos y que, cuando un JM es correcto, debe serlo, como mínimo, para toda una comunidad. Así, si resulta que el juicio «torturar a un ser humano es malo» es correcto, deberá serlo para toda una comunidad, o en otras palabras, todos los miembros de una comunidad deberán admitir que torturar a un ser humano es malo. El punto de Brink es que, para poder dar cuenta de este carácter objetivo de la moralidad, es necesario concebir a los JMs como creencias que describen estados de cosas independientes de los sujetos que sostienen esos JMs. Así, Brink termina inclinándose hacia la tesis de que los JMs son tan objetivos como lo son los enunciados propios de las ciencias empíricas. Pero una consecuencia inmediata de tal concepción, es que los JMs, contrario a lo que piensan los Internalistas, no son intrínsecamente motivacionales, esto es, no motivacionales en virtud de su contenido. Esto se debe a que, como vimos en el capítulo 1, de acuerdo con la Psicología Humeana, las creencias no pueden concebirse como estados mentales propiamente motivacionales. Ellas son incapaces de influir directamente en, o causar directamente, nuestras acciones. Las creencias sólo pueden influir en nuestras acciones en conjunción con un deseo o, en

general, un estado que sí sea intrínsecamente motivacional. Así, según Brink, a fin de dar cuenta de la objetividad del lenguaje y del pensamiento moral, es necesario rechazar el Internalismo. En el capítulo 3, ya rechazamos esta concepción particular de la objetividad y argumentamos que el Realismo Moral es insostenible. Sin embargo, los argumentos de Brink a favor del Externalismo son independientes de su Realismo Moral; por tal razón, ahora los consideraremos separadamente.

El punto de partida de Brink es la idea de que la moralidad tiene un carácter esencialmente práctico, es decir, está ligada inevitablemente a la acción. cualquier teoría metaética, incluso el Externalismo, debe estar en capacidad de explicar este carácter especial. Esto es, justamente, lo que hace tan atractivo al Internalismo. Sin embargo, dado su compromiso con el Realismo Moral, Brink considera que éste es falso. Así, a fin de sustentar su Realismo, él concentra sus esfuerzos en realizar dos tareas: 1) dar argumentos adicionales, independientes de consideraciones acerca del Realismo, en contra del Internalismo; y 2) mostrar cómo un modelo Externalista puede dar cuenta del carácter práctico de la moralidad.

Con respecto al primer punto, Brink sostiene que las afirmaciones de Hare acerca de los sociópatas o amorales, que involucran la noción de uso de conceptos morales en comas-invertidas, no toman realmente en serio las pretensiones de amoralidad. Según él:

El problema del Internalismo es que no toma suficientemente en serio el planteamiento del amoral. Independientemente de cuáles sean los méritos de esta estrategia Internalista como interpretación de Trasímaco o Uriah Heep, lo que plantea el amoral no depende necesariamente de que no se reconozca el uso entrecorillado de los términos. Podemos imaginar a alguien que considere ciertas pretensiones como pretensiones morales –y no simplemente como pretensiones morales convencionales– y que, no obstante, permanezca impasible. [...] pero efectivamente creemos que tal persona es posible y, si pensamos tomar seriamente el planteamiento del amoral, debemos intentar explicar por qué a este debería preocuparle la moralidad. (1986, p. 9)

El punto de Brink parece ser que argumentar del modo en que lo hace Hare es asumir por principio el Internalismo. En efecto, ¿por qué debemos asumir que el amoral o el sociópata no tienen plena comprensión de los términos morales? ¿por qué debemos pensar que el uso que tales personas hacen de los términos morales no es un uso legítimo, sino un uso en comas-invertidas? Las afirmaciones acerca del uso en comas-invertidas sólo pueden formularse como una *explicación* de por qué las personas, que no poseen comprensión de los términos morales, aún así, son capaces de comportarse lingüísticamente de un modo similar a como lo haría cualquier persona que sí los comprenda. Pero, para llegar a plantear tal explicación, antes es necesario ofrecer argumentos independientes de las consideraciones acerca del uso en comas-invertidas que muestren que, en efecto, tales personas no comprenden los términos morales. La ausencia de tales argumentos en los planteamientos de Hare, muestra que él simplemente está dando por supuesto que los amorales y los sociópatas no comprenden el lenguaje moral. Y esto es, precisamente, no tomar en serio las pretensiones del amoralista.

Puestas las cosas de este modo, es claro que lo que está en juego en la polémica entre Internalistas y Externalistas es, en últimas, la cuestión acerca de *en qué consiste comprender un término moral*. Los Internalistas, del estilo de Hare, piensan que la comprensión de los términos morales está constituida, al menos en parte, por la motivación moral. En contraste, los Externalistas, como Brink, consideran que el que un sujeto tenga un comportamiento lingüístico normal es un criterio suficiente para la atribución de comprensión del lenguaje moral; precisamente, por tal razón, Brink piensa que es posible atribuirle comprensión a los amorales y a los sociópatas, ya que ellos se



comportan lingüísticamente del mismo modo en que lo hacen las personas normales. Así, el punto de Brink es que Hare parte de una teoría acerca de la comprensión de los términos morales, que de manera inmediata implica el Internalismo. En otras palabras, al formular el argumento del uso en comas-invertidas, Hare ya da por sentado que no es posible que existan individuos amorales, en el sentido que los Externalistas le dan a este términos, esto es, individuos que comprenden plenamente el lenguaje moral y, a la vez, carecen por completo de la motivación para actuar de conformidad con sus JMs. En vista de esto, para que el Internalismo tenga algo de plausibilidad, sus defensores deben ofrecer argumentos a su favor, que sean independientes de las consideraciones acerca del uso en comas-invertidas. Teniendo en cuenta esto, Brink afirma que la posibilidad de que existan sujetos amorales es un reto que deben superar los Internalistas, ellos deben mostrar, o bien que realmente no existen tales sujetos, o bien que aquellos que consideramos amorales en realidad son personas incapaces de comprender los términos morales y de emitir, en sentido estricto JMs. A su modo de ver, éste es un reto que el Internalismo no puede superar; por tal razón, Brink concluye que necesariamente debemos inclinarnos hacia el Externalismo.

Hasta acá, hemos mostrado la manera en que Brink considera que el Realismo puede cumplir con la primera tarea que tiene, a saber, dar argumentos en contra del Internalismo, que sean independientes de las consideraciones acerca de la objetividad del lenguaje moral. Sin embargo, todavía queda por cumplir la segunda tarea, a saber, mostrar de qué modo el Externalismo puede dar cuenta del carácter práctico de la moralidad. ¿Cómo puede un modelo Externalista explicar el hecho de que intuitivamente parece haber una conexión entre nuestros JMs y nuestras motivaciones para actuar? ¿cómo puede el Externalismo explicar el hecho de que los JMs guían nuestra conducta? Sobre este punto, Brink afirma:

[...] el realismo moral es perfectamente compatible con el carácter práctico de la moralidad. Esto se debe a que el externalismo, a diferencia del internalismo, es el modo adecuado de representar el carácter práctico o de guía de la acción que reviste la moralidad. La racionalidad y la fuerza motivacional de las consideraciones morales dependen, como sostiene el Externalista, no sólo del concepto de moralidad sino (además) del contenido de la moralidad, los hechos sobre los agentes y una teoría sustantiva de las razones para la acción (1986, p. 17).

Según esto, para Brink, la explicación de cómo es posible la motivación moral depende, principalmente, de una teoría acerca de las razones para actuar. En particular, Brink se contempla dos posibilidades: el *egoísmo racional objetivista* y el *altruismo racional*. Según el egoísmo racional, “un agente tiene una razón para hacer x, cuando hacer x está en su propio interés” (1986, p. 12). Además, en la versión objetivista de dicha teoría “el interés del agente consiste en cuestiones como tener un cierto tipo de carácter, participar de cierto tipo de actividades y ejercitar ciertas capacidades (*Idem*) Por su parte, el altruismo racional “sostienen que alguien tiene una razón para hacer x solo si está en el interés de *alguien* que x se haga” (*Idem*) El punto de Brink, entonces, es que la motivación moral debe explicarse, como hemos dicho antes, a partir de hechos adicionales acerca de las razones para actuar de un agente; específicamente, a partir de un *deseo de hacer lo bueno* (sea lo que sea), el cual es independiente del contenido mismo de los JMs particulares. Qué tipo de deseo es éste, es algo que puede decirse a partir de alguna de estas teorías de las razones mencionadas. Así, por ejemplo, si se adopta el egoísmo racional; el agente desearía realizar cualquier acción que sea buena, debido a que la realización de ella contribuye a la satisfacción de su propio interés. Según el Externalismo, la conjunción entre un deseo como éste y, por ejemplo, el juicio de que “donar dinero para ayudar a los pobres es bueno” lleva a que el agente tenga la

motivación particular de donar dinero para ayudar a los pobres. De este modo, según Brink, es posible explicar el carácter práctico de la moralidad, sin comprometerse con la tesis Internalista de que las motivaciones son constitutivas del contenido de los JMs.

#### 4.4. La Defensa de Smith del Internalismo y el Fetichismo Moral

Michael Smith ha sido el primer defensor del Internalismo, en asumir el reto planteado por Brink. Smith reconoce que la aceptación del argumento del uso en comas-invertidas depende de la aceptación de una teoría acerca de la comprensión del lenguaje moral, en la que el comportamiento lingüístico no es un criterio suficiente para la atribución de competencia en el uso de dicho lenguaje. En la medida en que no hay ninguna razón por la que debamos exigirle al Externalista que acepte esta teoría sobre de la comprensión del lenguaje moral, no podemos usar el argumento del uso en comas-invertidas contra él, a fin de resolver la polémica. Por consiguiente, para mostrar que realmente no es posible que existan sujetos amorales, en el sentido que el Externalismo le da a esta expresión, es necesario dar un argumento adicional a favor de Internalismo (Cf. Smith 1994, pp. 68-71). El capítulo 3 de *The Moral Problem* está dedicado, justamente, a tratar de formular este argumento adicional.

La argumentación de Smith comienza con la consideración de que, tanto el Internalismo como el Externalismo deben dar cuenta de los *cambios en la motivación* de un agente. Alguien puede, por ejemplo, juzgar que debe votar por demócratas y, dado el carácter práctico de la moralidad, estar motivado a votar por demócratas. No obstante, al involucrarse en ciertas discusiones políticas, dicho agente puede convencerse de que su Juicio inicial estaba errado y, a partir de ese momento, comenzar a juzgar que en realidad debe votar por republicanos. Nuevamente, dado el carácter práctico de la moralidad, ahora comenzará a estar motivado a votar por republicanos. Tanto el Internalismo como el Externalismo deben poder explicar cómo son posibles estos cambios. Para los Internalistas, proveer tal explicación no representa mayores problemas. Dado que, según ellos, la conexión entre los JMs y las motivaciones es interna; ellos simplemente puede decir que, al cambiar el JM, necesariamente debe cambiar la motivación inherente a él. Según Smith, esta explicación, permite concebir a la motivación moral en términos de un deseo *de re*. Así, por ejemplo, si un agente *A* juzga que para él  $\phi$ ar es bueno; este Juicio, o bien *se identificará con*, o bien *producirá directamente*, un deseo de *A* cuya forma lógica puede expresarse del siguiente modo:

$$\exists x [(A \text{ Desea}(\text{realizar } x)) \wedge (x \text{ es bueno}) \wedge (x = \phi ar)]$$

Según este modelo, el cambio de motivación ocurre de manera directa en virtud del cambio de Juicio. Esto se debe a que, al cambiar el Juicio de *A* de que  $\phi$ ar es bueno, necesariamente se producirá un nuevo deseo con la misma forma; pero orientado, ya no a  $\phi$ ar, sino hacia la acción que ahora se juzga como buena. Por su parte, los Externalistas deben contar una historia un poco más compleja. Dado que, como vimos al considerar los planteamientos de Brink, ellos niegan que la motivación es inherente a los JMs; ellos deben explicar los cambios de motivación, apelando a un deseo general de hacer lo que es bueno. Según Smith, tal deseo debe ser un deseo *de dicto*, cuya forma lógica es probablemente la siguiente:

$$A \text{ Desea}[\exists x(x \text{ es bueno}) \supset (\text{realizar } x)]$$

Según este modelo, si un agente *A* juzga que  $\phi$ ar es bueno, su motivación será un deseo de  $\phi$ ar *derivado* de su juicio y del deseo general *de dicto* de hacer cualquier acción que se evalué como buena. De este modo, el cambio de motivación se producirá, debido a que el cambio de Juicio permite derivar un deseo particular diferente. Vemos, así, que tanto el Internalismo como el Externalismo tienen recursos para explicar satisfactoriamente los cambios de motivación.

El segundo paso en la argumentación de Smith consiste en resaltar que, a partir de lo anterior, es evidente que cada modelo da una imagen diferente de *lo que es ser un agente moral*, esto es, de lo que es ser un hombre bueno. Según el modelo Internalista, un hombre bueno siempre se encuentra motivado a actuar *en virtud de las acciones mismas*; esto quiere decir que él desea realizar las acciones buenas, porque encuentra en ellas algo que le resulta *por sí mismo deseable*. Esto se manifiesta claramente en el hecho de que sus motivaciones tienen la forma de un deseo *de re*. En contraste, de acuerdo con el modelo Externalista, un hombre bueno es aquel que posee un deseo general *de dicto* de realizar cualquier acción que sea buena. Según esto, la diferencia entre los sociópatas y amorales, por una parte, y los hombres buenos, por la otra, radica, no en cuestiones relativas a la comprensión que unos y otros tienen del lenguaje moral; sino en que, los primeros, pero no los segundos, carecen de tal deseo. Smith considera que esta concepción Externalista del hombre bueno es problemática. En efecto, de acuerdo con ella, los motivos morales de todo hombre bueno no dependen de rasgos que las acciones poseen en sí mismas, es decir, no se producen por la apreciación de que las acciones que se consideran buenas son en sí mismas deseables. Los motivos morales son, para el agente, simples *medios*, que le permiten satisfacer otro deseo, a saber, el deseo general *de dicto* de realizar cualquier acción que sea buena. En otras palabras, la concepción Externalista del hombre bueno lo presenta como alguien para quien los motivos morales tienen un carácter puramente *instrumental*. Según Smith, esta visión puede ser considerada una forma de lo que él denomina “fetichismo moral”, algo que, claramente, va en contra de la mayor parte de nuestras intuiciones acerca de lo que es ser un hombre bueno (Cf. 1994, pp. 71 – 6).

Para Smith, estas afirmaciones constituyen el argumento adicional que Brink le exige a los defensores del Internalismo. El Externalismo, según esto, es una doctrina insostenible ya que, en cualquier de sus formas, termina cayendo en el fetichismo moral y negando muchas de nuestras ideas generales acerca de la moralidad.

## BIBLIOGRAFÍA

- Ayer, A.J., 1952, *Language, Truth, and Logic*, New York: Dover. Edición española: Ayer, A.J., 1975, *Lenguaje, Verdad y Lógica*, Buenos Aires: Eudeba. Referencias a la edición española.
- Bermúdez, José Luis, 1996, "The moral significance of birth", *Ethics*, Vol. 106, No. 2, pp. 378-403
- Brink, David O., 1986, "Externalist moral realism", *The Southern Journal of Philosophy*, Supplement, pp. 23-42. Traducción inédita al español de Max Kölbel: "El realismo moral externalista", pp. 1-17. Referencias a la traducción al español.
- \_\_\_\_\_ 1989, *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- \_\_\_\_\_ 1997, "Moral motivation", *Ethics*, Vol. 108 (1), pp. 4-32.
- Cohn, J.; & Tronick, E.; 1987, "Mother-infant face-to-face interaction: The sequencing of dyadic states at 3, 6, and 9 months", *Developmental Psychology*, 23, pp. 68-77.
- \_\_\_\_\_ 1988, "Mother-infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior", *Developmental Psychology*, 24, pp. 386-392.
- Cussins, A., 1994, "La Construcción Conexionista de los Conceptos", en: M. Boden (ed.), *La Filosofía de la Inteligencia Artificial*. México: FCE. Trad.: Guillermina Feher de la Torre, pp. 409-88.
- \_\_\_\_\_ 2002, "Experience, Thought, and Activity". Sin publicar.
- \_\_\_\_\_ 2003, "Content, Conceptual Content, and Nonconceptual Content", en: Gunther, Y.H. (ed.), *Essays on Nonconceptual Content*. Cambridge & London: The MIT Press, pp. 133 – 64.
- Darwall, Stephen; Gibbard, Allan; & Railton, Peter; 1992, "Toward Fin de siecle Ethics: Some Trends", *The Philosophical Review*, Vol. 101, No. 1, Philosophy in Review: Essays on Contemporary Philosophy, pp. 115-189.
- Davidson, Donald, 1970, "How is weakness of the will posible?", en: Davidson 1980, pp. 21-42.
- \_\_\_\_\_ 1980, *Essays on Actions and Events*, Oxford: Clarendon Press.
- Dreier, James, 2000, "Dispositions and fetishes: externalist models of moral motivation", *Philosophy and Phenomenological Research*, Vol. LXI, No. 3, pp. 619-38.
- Dummett, Michael, 1978, *Truth and Other Enigmas*, Cambridge & London: Harvard University Press.

- \_\_\_\_\_ 1978a, "Truth", en: 1978, pp: 1-24.
- \_\_\_\_\_ 1978b, "Wang's Paradox", en: 1978, pp.: 248-68.
- \_\_\_\_\_ 1991 *The Logical Basis of Metaphysics*, Cambridge: Harvard University Press.
- \_\_\_\_\_ 1993, *The Seas of Language*, Oxford : Clarendon Press.
- \_\_\_\_\_ 1993a, "Realism and Anti-Realism", en 1993a, pp. 462 – 78.
- \_\_\_\_\_ 2006, *Thought and Reality*, Oxford: Oxford University Press.
- Frankena, W.K., 1939, "The naturalistic fallacy", *Mind*, New Series, Vol. 48, No, 192, pp. 464-477.
- Gallagher, Shaun, 1996, "The moral significance of primitive self-consciousness: A response to Bermúdez", *Ethics*, Vol. 107, No. 1, pp. 129-140
- \_\_\_\_\_ 2001, "The practice of mind. Theory, simulation or primary interaction?", *Journal of Consciousness Studies*, 8, No. 5–7, 2001, pp. 83–108
- Gallagher, Shaun; & Zahavi, Dan, 2008, *The Phenomenological Mind*, London & New York: Rodledge.
- Gibson, Eleanor J.; & Pick, Anne D., 2000, *An Ecological Approach to Perceptual Learning and Development*, Oxford: Oxford University Press.
- Gibson, J.J., 1986, *The Ecological Approach to Visual Perception*, New Jersey & London: Lawrence Erlbaum Associates, Publishers.
- Gibbard, A., 1990, *Wise Choices, Apt Feelings*, Cambridge: Harvard University Press.
- Gunther, Y.H., 2003, "Emotion and Force", en: Gunther, Y.H. (ed.), *Essays on Nonconceptual Content*. Cambridge & London: The MIT Press, pp. 133 – 64.
- Hare, R.M., 1952, *The Language of Morals*, Oxford: Oxford University Press.
- \_\_\_\_\_ 1997, *Sorting out Ethics*, Oxford: Clarendon Press. Edición española: Hare, R.M., 1999, *Ordenando la Ética*, Barcelona: Ariel, trad.: Joan Vergés Gifra. Referencias a la edición española.
- Harman, G., 2000, "Moral Relativism Defended", en: *Explaining Value and Other Essays in Moral Philosophy*, Oxford: Oxford University Press, pp. 3-19.
- Kessen, W.; Levine, J.; & Wendrich, K.A.; 1979, "The initiation of pitch in infants", *Infant Behavior and Development*, 2, pp. 93-100.
- Legerstee, M., 1990, "Infants use multimodal information to imitate speech sounds", *Infant Behavior and Development*, 13, pp. 343-354.
- \_\_\_\_\_ 1991, "The role of person and object in eliciting early imitation", *Journal of Experimental Child Psychology*, 51, pp. 423-433.
- Mackie, J.L., 1977, *Ethics: Inventing the Right and Wrong*, London: Pelican Books. Edición española: Mackie, J.L., 2000, *Ética: Inventando el Bien y el Mal*, Barcelona: Gedisa, trad.: Tomás Fernández Aúz. Referencias a la edición española.
- McDowell, Michael, 1998, *Mind, Value, and Reality*, Cambridge and London: Harvard University Press.
- \_\_\_\_\_ 1998a, "Are moral requirements hypothetical imperatives?", en: McDowell, John 1998, pp. 77-94.

- \_\_\_\_\_ 1998b "Non-Cognitivism and Rule-Following", en: McDowell, John 1998, pp. 198 – 218.
- Meltzoff, A.; & Moore, M., 1983, "Newborn infants imitate adult facial gestures", *Child Development*, 54, pp. 702-709.
- Miller, Alexander, 2003, *An Introduction to Contemporary Metaethics*, Cambridge: Polity.
- Moore, G.E., 1903, *Principia Ethica*, Cambridge: Cambridge University Press. Edición española: Moore, G.E., 2002, *Principia Ethica*, Barcelona: Crítica, trad.: María Vásquez Guisán. Referencias a la edición española.
- Peacocke, C., 1983, *Sense and Content: Experience, Thought, and Theirs Relations*, Oxford: Clarendon Press.
- Platts, Mark, 1979, *Ways of Meaning*, London: Routledge and Kegan Paul
- Prinz, J., 2006, "The Emotional Basis of Moral Judgement", *Philosophical Explorations* Vol. 29, No. 2.
- Railton, P., 1986, "Moral Realism", *Philosophical Review*, pp. 163 – 207.
- Roskies, Adina, 2003, "Are ethical judgement intrinsically motivational? Lessons from 'acquired sociopathy'", *Philosophical Psychology*, Vol. 16, No. 1, pp. 51-66
- Shafer-Landau, R., 1994, "Ethical Disagreement, Ethical Objectivism and Moral Indeterminacy", *Philosophy and Phenomenological Research*, Vol. LIV, No. 2, pp.: 331-44.
- Sinnott-Armstrong, Walter (ed.), 2008, *Moral Psychology, Vol 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, Boston: The MIT Press.
- Smith, Michael, 1994, *The Moral Problem*, Oxford: Blackwell
- \_\_\_\_\_ 2008, "The Truth About Internalism", en: Sinnott-Armstrong 2008, pp. 207-216.
- Stevenson, Charles Leslie, 1937, "The Emotive Meaning of Ethical Terms", *Mind*, New Series, Vol. 46, No. 181, pp. 14-31. Traducción al español "El significado emotivo de los términos éticos" en: Ayer, A.J. (Comp.), 1993, *El Positivismo Lógico*, México: FCE, trad.: L. Aldama, U. Frisch, C.N. Molina, F.M. Torner y R. Ruiz, pp. 269-286. Referencias a la traducción al español
- Stevenson, C.L., 1944, *Ethics and Language*, New Haven, Conn.: Yale University Press.
- Sturgeon, N., 1988, "Moral Explanations", en: J. Sayre-McCord (Ed.), *Essays on Moral Realism*, Ithaca & London: Cornell University Press, pp. 229 – 55.
- Vintner, A., 1986, "The role of movement in eliciting early imitations", *Child Development*, 57, pp. 66-71.
- Walker-Andrews, A.S.; & Lennon, E., 1991, "Infants' discrimination of vocal expressions: Contributions of auditory and visual information", *Infant Behavior and Development*, 14, pp. 131-142.
- Wedgwood, Ralph, 2007, *The Nature of Normativity*, Oxford: Oxford University Press
- Wright, C., 1992, *Truth and Objectivity*, Cambridge & London: Harvard University Press