

LA TESIS  
EMOCIONALISTA  
SOBRE LOS  
JUICIOS DE  
VALOR

---

JULIÁN CUBILLOS

---

UNIVERSIDAD NACIONAL DE COLOMBIA

FACULTAD DE CIENCIAS HUMANAS

DEPARTAMENTO DE FILOSOFÍA

MAESTRÍA EN FILOSOFÍA

«LA TESIS EMOCIONALISTA SOBRE LOS  
JUICIOS DE VALOR»

TESIS DE GRADO PARA OPTAR AL TÍTULO DE:

MAGÍSTER EN FILOSOFÍA

ESTUDIANTE: JULIÁN ALBERTO CUBILLOS OCAMPO

DIRECTORA: DRA. ÁNGELA URIBE BOTERO

BOGOTÁ D.C., CIUDAD UNIVERSITARIA

MARZO DE 2010

## AGRADECIMIENTOS

A mi familia, a la Bebé y a mis amigos, sin cuyo amor y motivación no podría emprender nada.

A la profesora Ángela Uribe por haber despertado en sus enriquecedoras clases mi interés por la filosofía moral y por la valiosísima ayuda que me dio como directora de esta tesis.

Al profesor Jesse J. Prinz, pues si bien lamento no haber tenido a tiempo una traducción de mi tesis al inglés, con el objeto de recibir su valoración, agradezco mucho su amable motivación en el desarrollo de la misma y su gentil aprobación para publicar mi traducción al español de su artículo “Emotion and Aesthetic Value”, en el que se apoya el tercer capítulo de mi trabajo.

A la profesora Catherine Z. Elgin y a su muy querida alumna Remei Capdevila por su inmenso apoyo en mi comprensión de la filosofía del arte. A Gabriel Jiménez y a Andrés Obando por sus valiosas charlas sobre filosofía moral. A los profesores, compañeros y alumnos que me han acompañado en estos primeros pasos en el estudio de la moral y la estética. A todos y todas, mis más sinceros agradecimientos.

---

# CONTENIDO

INTRODUCCIÓN .....	<i>i</i>
--------------------	----------

## CAPÍTULO 1

### PRESUPUESTOS NATURALISTAS DE LA TESIS EMOCIONALISTA

1.1 El naturalismo moral.....	2
1.1.1 El naturalismo moral frente a la Ley de Hume.....	7
1.1.2 El naturalismo moral y la ‘falacia naturalista’.....	14
1.2 El enfoque naturalista.....	23

## CAPÍTULO 2

### LA TESIS EMOCIONALISTA EN LA EXPLICACIÓN COGNITIVA DEL JUICIO MORAL

2.1 El emocionalismo moral .....	30
2.1.1 El modelo “intuicionista social” de Jonathan Haidt .....	30
2.1.2 El emocionalismo de Jesse J. Prinz .....	38
2.2 El innatismo moral .....	52
2.2.1 La competencia moral en Susan Dwyer .....	53
2.2.2 El órgano moral en Marc Hauser .....	56
2.3 El híbrido emocionalismo-racionalismo moral .....	66
2.3.1 El proceso dual de la moral en Joshua Greene .....	67
2.4 ¿Un diálogo de sordos? .....	70

### CAPÍTULO 3

#### LA TESIS EMOCIONALISTA EN LA EXPLICACIÓN COGNITIVA DEL JUICIO DE GUSTO ESTÉTICO

3.1	Arte y emoción .....	75
3.1.1	Una teoría afectiva de la apreciación estética .....	76
3.2	Arte y conocimiento .....	85
3.2.1	El lugar del entendimiento en la apreciación del arte: un argumento técnico.....	85
3.2.2	El lugar del entendimiento en la apreciación del arte: un argumento intuitivo.....	100
3.3	Un arte apasionado pero informado .....	102
	CONCLUSIONES .....	106
	APÉNDICE I .....	113
	APÉNDICE II .....	117
	BIBLIOGRAFÍA .....	121

---

# INTRODUCCIÓN

En líneas generales, la que aquí propongo denominar la ‘tesis emocionalista sobre los juicios de valor’ (o, para abreviar, ‘tesis emocionalista’) puede ser caracterizada como la conjunción de dos afirmaciones: *i.* “(...) las emociones no están simplemente relacionadas con los juicios morales sino que ellas también son, en algún sentido, tanto necesarias como suficientes para dichos juicios” (Prinz, 2006, 29), y *ii.* “(...) cuando apreciamos una obra de arte, la apreciación consiste en una respuesta emocional” (Prinz, 2007b, 1). La tesis emocionalista sobre los juicios de valor está encaminada, así, a defender que pese a las diferencias importantes que hay entre los ámbitos de la valoración moral y la valoración estética, ambos tienen un fundamento afectivo y que, en consecuencia, podemos defender que estos se construyen emocionalmente.

Esta tesis ha cobrado su mayor fuerza y poder explicativo en la filosofía de Jesse J. Prinz, puesto que, como veremos en este trabajo, es él quien la ha expuesto de la manera más sistemática, completa, coherente y fundamentada en el panorama de la filosofía actual. Este es el resultado del desarrollo, *in extenso*, que Prinz hace de los planteamientos de Hume con respecto al fundamento emocional de los juicios de valor, complementando el trabajo de este último con la evidencia que arroja la psicología empírica contemporánea. Se trata, así, de una postura humeana corregida en algunos

aspectos y fortalecida con los hallazgos de la psicología, neurociencia cognitiva, investigación sobre psicopatología y observaciones antropológicas.

Por esta razón, permítaseme comenzar exponiendo la manera en que Prinz considera que su tesis se erige a partir de su interpretación de la obra humeana; exposición que, dicho sea de paso, me servirá para presentar el trabajo general de la obra de Prinz. Antes, sin embargo, es de anotar que —atendiendo a la forma en que el autor ha presentado su trabajo— la segunda parte de la tesis de Prinz resulta realmente una extensión de la primera. Con esto me refiero a que el autor ha dedicado todos sus libros a fundamentar una construcción emocional de la moral y tan solo en un par de artículos ha sostenido que, *mutatis mutandis*, gran parte del fundamento emocional que sostiene esta construcción bien se puede aplicar a una construcción de la valoración estética. Esto explica por qué en esta presentación la articulación de Hume con la obra de Prinz se centra más en el juicio moral que en el estético.

Hecha esta aclaración, comenzaré diciendo, entonces, que así como Hume dividió su *Tratado de la naturaleza humana* en tres libros: “Sobre el entendimiento”, “Sobre las pasiones” y “Sobre la moral”, Prinz interpreta esta obra a partir de un eje conductor: el fundamento emocional. De esta manera, el autor señala que Hume desarrolla una teoría de los conceptos (o “ideas”) en el primer libro y una teoría de las emociones en el segundo libro, y que luego integra estos dos en su tercer libro, defendiendo así que nuestros conceptos morales tienen un fundamento emocional. Además, para Prinz, Hume unifica el proyecto mediante su defensa del empirismo, en tanto que su teoría de los conceptos está construida sobre la premisa de que estos conceptos o ideas son copias que se suplen de impresiones sensoriales y su teoría de las emociones está diseñada para ser compatible con su punto de vista empirista —las emociones quedan definidas, así, como impresiones de impresiones—. De acuerdo con

esto, concluye Prinz, la teoría moral de Hume también es empirista (cf. Prinz, 2007a, vii).

Sin embargo, Prinz señala que los conceptos morales podrían constituir un contraejemplo para una postura empirista que, como la de Hume, conciba los conceptos como copias que se suplen de impresiones sensoriales. Esto en razón de que, por ejemplo, no hay una imagen de la virtud, ni un gusto de lo bueno y tampoco un olor de lo malo. De aquí que, para Prinz, Hume apele a los sentimientos para resolver este problema, defendiendo que si bien todos los conceptos están basados en impresiones, el concepto de lo bueno se basa en un sentimiento de aprobación y el concepto de lo malo se basa en un sentimiento de desaprobación. Así, aunque el conjunto de las virtudes no tiene una apariencia en común, podríamos decir que las cosas buenas nos hacen sentir justamente bien; de manera similar, y aunque sería imposible trazar el conjunto total de los vicios, también podríamos decir que de cada instancia de estos se deduce un palpable sentimiento de culpa o remordimiento. Es pues de este modo como Prinz ve que el *Tratado* de Hume tiene una estructura coherente y que la teoría moral culminante puede ser leída como la resolución de un aparente contraejemplo a su teoría de los conceptos. En consecuencia, para el filósofo norteamericano, sin importar en dónde se ponga el énfasis, la teoría de los conceptos de Hume y su teoría de la moral van de la mano, y la emotividad es la articulación de estas dos (cf. *ibíd.*).

A partir de esta interpretación, en su último libro —*The Emotional Construction of Morals*—, Prinz defiende una teoría sentimental de la moralidad que se construye sobre las ideas desarrolladas por Hume y algunos de sus contemporáneos. El autor confiesa que, pese a tener algunas diferencias con Hume, la estructura básica de su teoría también es humeana y que, a este respecto, sus propuestas son pies de página al Libro



III del *Tratado* (cf. *ibíd.*). Prinz también es el autor de *Furnishing the Mind* (Prinz, 2002), en donde defendió una teoría empirista de los conceptos, y de un tercer libro, titulado *Gut Reactions* (Prinz, 2004a), en donde defendió una teoría empirista de las emociones. Y esto es pertinente en tanto que, en su último libro, *The Emotional Construction of Morals* (Prinz, 2007a), Prinz confiesa estar completando una trilogía que corre paralela a la estructura del *Tratado* de Hume. De esta manera, aunque cada uno de los trabajos sea autónomo, ellos van de la mano justamente en la forma en que el *Tratado* de Hume va de la mano: constituyen una construcción empirista de la emotividad como fundamento de la moral, esto es, una construcción emocional de la moralidad.

Con este proyecto, los propósitos principales de Prinz son tres. En primer lugar, él trata de proveer un soporte empírico para esta construcción emocional de la moralidad. En segundo lugar, Prinz intenta complementar la teoría de Hume, incluyendo la defensa de que los sentimientos soportan nuestros juicios morales y una defensa de la ontología que resulta de tomar seriamente un punto de vista sentimentalista. Por último, el autor quiere mostrar que esta aproximación nos lleva al relativismo moral, relativismo al que Hume se resistió, pero que —desde el punto de vista de Prinz—, era indefendible<sup>1</sup>.

Es así como Prinz indaga en los orígenes de nuestros sentimientos morales y sugiere que la aproximación genealógica de Nietzsche a la moralidad tiene mucho para

---

<sup>1</sup> Esto porque Hume sugiere que los desacuerdos morales deben estar basados en desacuerdos sobre hechos no-morales, ya que todas las personas, por naturaleza, tenemos los mismos valores morales fundamentales —valores que Prinz interpreta como aquellas normas fundamentales [*grounding norms-norms*] que, psicológicamente, no dependen de la apelación a los demás—. Esta afirmación, según Prinz, presupone que los valores fundamentales se derivan de la naturaleza humana. Pero, aun aceptando este punto —con el que, sin embargo, Prinz tampoco estará de acuerdo (cf. 2007a, cap. 7)—, no se sigue que estas normas fundamentales sean fijas o rígidas. Para Prinz, la educación moral se da mediante el condicionamiento emocional, condicionamiento que puede alterar disposiciones afectivas establecidas con anterioridad. De manera similar a como podemos adquirir nuevos caprichos, miedos y fantasías mediante el condicionamiento, también podemos adquirir nuevas normas fundamentales. Así, por ejemplo, dos personas que hayan sido condicionadas de manera diferente durante su desarrollo tendrán diferentes valores morales fundamentales (cf. 2007a, 194-95).

decir aquí. Nietzsche quiso desestabilizar nuestros valores mediante la exposición de su pasado y, en esta crítica, una de las grandes virtudes de su proyecto es la de mostrar que las sociedades crean diferentes sistemas morales debido a sus condiciones históricas diferentes. A partir de esta intuición Prinz se propone, por un lado, demostrar que el método genealógico se puede utilizar de manera eficaz para investigar el origen de los valores, puesto que ayuda a confirmar que algunas convicciones morales son producto de la historia social; y, por otro lado, que la genealogía no apoya por completo el escepticismo sobre los valores morales. La historia resultante es, entonces, medio humeana y medio nietzscheana, pero con la parte nietzscheana Prinz busca armonizar la parte humeana. Finalmente, Prinz señala la manera en que su trabajo puede verse como una continuación de la tradición de Westermarck, en cuanto que fue él quien —hace ya casi un siglo— reconoció la relación entre el sentimentalismo y el relativismo, y reconoció el valor de la antropología y la historia en investigaciones morales.

Considero que esta presentación del trabajo de Prinz es pertinente para apoyar las razones por las cuales encuentro que su postura sistematiza muy bien el emocionalismo. Sin embargo y pese al gran atractivo de esta teoría, encuentro que la tesis emocionalista, aun cuando se presente *prima facie* como la mejor explicación, deja algunos vacíos que bien podrían ser llenados complementando esta postura con una aproximación cognitiva racional. Por esta razón, mi propósito principal en este trabajo es el de defender que aunque la tesis emocionalista sobre los juicios de valor parece suficiente para dar cuenta de los procesos mediante los cuales surgen estos juicios, esta tesis no logra ofrecer una adecuada y completa comprensión de los mismos. Así, el primer componente de la tesis emocionalista, como pretendo mostrar aquí, incurre en un radicalismo unilateral, en tanto que deja por fuera el no menor papel que cumple el razonamiento en la producción de los juicios morales. El segundo componente, por su

parte, malentendiendo la manera en que funcionan las obras de arte y sugiere, así, una comprensión sesgada de las mismas, al subvalorar el papel que cumple el entendimiento en la apreciación de una obra de arte en cuanto obra.

Para este efecto, he dividido este trabajo en tres partes. En primer lugar y a partir de los planteamientos de Prinz en su libro *The Emotional Construction of Morals*, trataré de explicar cuáles son los presupuestos naturalistas en los que se enmarca el proyecto emocionalista. Con esto, mi idea es resaltar las virtudes de una naturalización de la moral, de acuerdo con la cual los procesos que determinan esta última pueden ser completamente descriptibles.

En el segundo capítulo, mi idea es examinar el lugar del componente moral de la tesis emocionalista en el debate cognitivo actual sobre la naturaleza de los juicios morales. El problema subyacente a esta indagación, como mostraré allí, es el que parte de la cuestión acerca de si podemos afirmar que los juicios morales están constituidos principalmente sobre la base de la emoción, o si son primariamente producto de la razón. Yo creo que la cuestión, planteada así en términos de una disyunción excluyente, no hace justicia a la posibilidad de que ambos componentes aporten en la producción de los juicios morales. Así, para mostrar que esta posibilidad es igualmente viable y en contraste con lo que sostiene la tesis emocionalista de Prinz, me propongo hacer fuerte la idea de Joshua Greene —expuesta, principalmente, en su artículo “An fMRI Investigation of Emotional Engagement in Moral Judgment” (cf. Greene et al., 2001)—, según la cual tanto la razón como la emoción cumplen un papel igualmente importante en la producción de los juicios morales.

Finalmente, en el tercer capítulo y mediante el examen de los argumentos expuestos por Prinz en su artículo “Emotion and Aesthetic Value” (cf. Prinz, 2007b),

intentaré mostrar la insuficiencia de la segunda parte de la tesis emocionalista, señalando cómo esta tesis malentiende la manera en que funcionan las obras de arte. Mi punto principal allí consistirá en defender que las obras de arte funcionan, en esencia, como una forma de ‘expresión’, de ‘expresión metafórica’, en particular; y de aquí que una respuesta emocional no sea suficiente para apreciar estéticamente una obra de arte, ya que la apreciación misma requiere, además, de un entendimiento mínimo de la manera en que funciona la expresión metafórica, esto es, entender el sentido en que la obra de arte funciona como obra de arte. Para este propósito, me apoyaré en los aportes que Nelson Goodman —en su libro *Los lenguajes del arte* (cf. Goodman, 1976)— y Catherine Z. Elgin —en *With Reference to Reference* (cf. Elgin, 1983)— han hecho sobre el problema que subyace a nuestro entendimiento y subsiguiente apreciación de las obras de arte, esto es, el problema de la referencia en las artes y lenguajes no verbales.

---

# 1. PRESUPUESTOS NATURALISTAS DE LA TESIS EMOCIONALISTA

Uno de los preceptos fundamentales que debe tener una filosofía propiamente cognitiva, si es que ha de tener alguno, es la idea de perseguir las explicaciones acerca de los hechos naturales con medios igualmente naturales. En los ámbitos de la moral y de la estética no podemos aceptar explicaciones de tipo supra-natural, por la sencilla razón de que estos ámbitos se derivan de nosotros y están, en consecuencia, anclados en la naturaleza humana y no en un reino espiritual. El método del análisis conceptual, por ejemplo, no puede constituir un método supra-natural para el descubrimiento de verdades supra-naturales, puesto que los conceptos mismos son entidades naturales y ellos pueden ser investigados usando procesos naturales.

A partir de estas acertadas intuiciones, que leyendo entre líneas pueden ser extractadas de la filosofía de Prinz, considero que es preciso determinar, por un lado, el sentido de naturalismo que se encuentra a la base de estas afirmaciones y, por otro lado, hacer una defensa de este a la luz del ámbito que nos concierne: los juicios de valor. Esto en razón de que, para tomar un caso particular, como lo es el de la moral, es bien sabido que las críticas a las propuestas naturalistas contemporáneas de la ética suelen apoyarse en la distinción ‘es/debe’ o ‘Ley de Hume’, expuesta por David Hume

(cf. Hume, 1981, Libro III, parte ii, sección 1), y en las objeciones de G. E. Moore (cf. Moore, 1997), conocidas como la ‘falacia naturalista’ y el argumento de ‘la pregunta abierta’.

Sin entrar aún a exponer en qué consisten estos argumentos, debo decir que su influencia ha sido tan notable en el ámbito de la ética contemporánea, que en mi opinión no se puede dar una base teórica legítima a las propuestas naturalistas sin ofrecer, primero, una solución naturalista a los problemas que señalan estos argumentos. Por esta razón, permítaseme apelar aquí a la exposición del naturalismo de Jesse J. Prinz (cf. Prinz, 2007a), postura que, según creo, cumple muy bien con el doble propósito de exponer una adecuada propuesta naturalista y de justificarla como forma de aproximación a los juicios de valor.

## 1.1 EL NATURALISMO MORAL

Prinz no pretende dar una definición última de un término tan ambiguo como lo es el ‘naturalismo’, más bien asume una tarea mucho más modesta y práctica: señalar en qué consiste el tipo de naturalismo con el que se compromete una explicación naturalista de la moral. Para él, el naturalismo es un conjunto de tesis que son contrarias a aquello que podemos denominar un supra-naturalismo —doctrina que busca fundamentaciones últimas en una especie de realidad que trasciende a los hechos— (cf. 2007a, 2).

Estas tesis, de acuerdo con Prinz, son cuatro: naturalismo metafísico, explicativo (o descriptivo), metodológico y transformativo. Para abreviar, he considerado pertinente sintetizar la tesis básica de cada una de estas clases de naturalismo en la siguiente tabla (v. Tabla 1):

	CLASES DE NATURALISMO	DESCRIPCIÓN
<b>NATURALISMO</b>	1. Naturalismo Metafísico	Se trata del punto de vista según el cual el mundo está limitado por postulados y leyes de las ciencias naturales. Nada puede existir sin que viole estas leyes y todas las entidades que existen deben, de alguna manera, estar compuestas de las entidades que contemplan nuestras mejores teorías científicas (cf. Prinz, 2007a, 2).
	2. Naturalismo explicativo o descriptivo	Si cada cosa que existe está compuesta de una sustancia natural y regida por ley natural, entonces cada cosa que no esté descrita en el lenguaje de una ciencia natural debe, en última instancia, ser descriptible en tales términos (cf. <i>ibíd.</i> ).
	3. Naturalismo metodológico	Si todos los hechos son, en algún sentido, hechos naturales (de acuerdo con el naturalismo metafísico), entonces los métodos mediante los cuales investigamos los hechos deben ser accesibles a la investigación de hechos naturales (cf. Prinz, 2007a, 3).
	4. Naturalismo transformativo	Se trata del punto de vista acerca de cómo cambiamos nuestras opiniones: estamos siempre operando desde adentro de las teorías actuales del mundo. Al hacer revisiones teóricas, no podemos salir de nuestras teorías y adoptar instancias trascendentales. Hacer eso sería suponer que tenemos una forma de pensar acerca del mundo que es independiente de nuestras teorías del mundo. Si las teorías del mundo provienen todas de nuestras creencias, entonces ninguna instancia trascendental es posible (cf. <i>ibíd.</i> ).

Tabla 1.

El autor señala que si estas clases de naturalismo son correctas, entonces los hechos morales son hechos naturales o no son hechos, puesto que los hechos naturales deben ser consistentes con dichas clases. Él hace una serie de aclaraciones sobre cada una de las tesis en cuestión (cf. Prinz, 2007a, 2-3); aclaraciones que no sobra anotar aquí porque, dicho sea de paso, nos servirán para sortear algunas dificultades que, a manera de objeción al naturalismo, podrían afectar también a una construcción naturalista de la moral —en particular, la objeción de que el naturalismo incurre en un reduccionismo fisicalista—.

Así pues, con respecto al naturalismo metafísico, Prinz aclara que se trata de una tesis metafísica y que, en consecuencia, concierne a la naturaleza fundamental de la realidad; de esta manera, el naturalismo metafísico constituye entonces el punto de partida de las otras tres tesis y funciona, así, como el antecedente de los condicionales en los que estas se sostienen. En cuanto al naturalismo descriptivo, Prinz sostiene que no se trata de un reduccionismo, en el sentido fuerte de esta palabra. Los reduccionistas fuertes afirman que la relación entre las ciencias naturales y “dominios de alto nivel” es deductiva y que, en consecuencia, estamos en la capacidad de deducir hechos de “alto nivel” de substratos de “bajo nivel”. Los anti-reduccionistas niegan esto. Ellos piensan, por ejemplo, que hay leyes o generalizaciones de “alto nivel” que bien podrían ser implementadas en un rango de maneras abiertamente-limitado y, así, regularidades obtenidas a un bajo nivel podrían perderse en dichas generalizaciones (cf. Prinz, 2007a, 2).

Prinz sostiene que la explicación naturalista puede ser anti-reduccionista, en tanto que esta no necesita defender que las explicaciones de bajo nivel sean las únicas explicaciones; su argumento radica en que hay algún género de correspondencia sistemática entre niveles y, así, según él: “Se debe estar en la capacidad de mapear cualquier entidad de alto nivel con una de bajo nivel, y se debe estar en la capacidad de explicar cualquier instanciación de cualquier generalización de alto nivel apelando a los caracteres de bajo nivel utilizados en estas generalizaciones” (ibíd.).

Con el tercer tipo de naturalismo, el metodológico, Prinz busca aclarar en qué sentido el análisis conceptual no puede constituir un método supra-natural para el descubrimiento de verdades supra-naturales —porque los conceptos en sí mismos son entidades naturales y ellos pueden ser investigados usando procesos naturales—. De este modo, si bien el análisis conceptual es, como todas las herramientas legítimas de



investigación, un método empírico, como tal no es especialmente poderoso. Esto en razón de que el análisis conceptual procede mediante el acceso en primera persona, o mediante introspección, a estructuras psicológicas; sin embargo, la introspección está propensa al error y es metodológicamente restringida, en tanto que se asocia con la obtención de conclusiones que usan a un sujeto particular (uno mismo). Por esta razón Prinz señala que:

[...] nosotros investigamos los conceptos usando herramientas de la ciencia social, puesto que si los conceptos son entidades naturales, entonces ellos surgen en formas naturales. Así, por ejemplo, los conceptos pueden ser adquiridos mediante la experiencia y ellos pueden ser revisados mediante la experiencia; ellos no tienen un estatus especial al momento de revelar hechos acerca del mundo (2007a, 3).

Por último, Prinz asocia el anterior tipo de naturalismo con W.V.O. Quine, quien propende porque la investigación del conocimiento pueda ser adelantada usando los recursos de las ciencias naturales. Así, de acuerdo con Prinz, Quine arguye que todas las afirmaciones están sujetas a revisiones empíricas. A partir de esta postura, el naturalismo metodológico señala entonces que estamos siempre operando desde adentro de las teorías actuales del mundo, puesto que al hacer revisiones teóricas no podemos salir de nuestras teorías y adoptar instancias trascendentales. Hacer eso sería suponer que tenemos una forma de pensar acerca del mundo que es independiente de nuestras teorías del mundo. Si las teorías del mundo provienen todas de nuestras creencias, entonces ninguna instancia trascendental es posible. Por tanto, el resultado de la combinación entre el naturalismo metodológico y las restricciones de Quine a las investigaciones del conocimiento, se puede llamar naturalismo transformativo, puesto que es la tesis acerca de la manera en que cambiamos nuestros puntos de vista sobre la realidad (cf. *ibíd.*).

Para sintetizar, pienso que la siguiente figura podría ayudar a obtener una mejor apreciación de estas cuatro consideraciones (v. Figura 1):

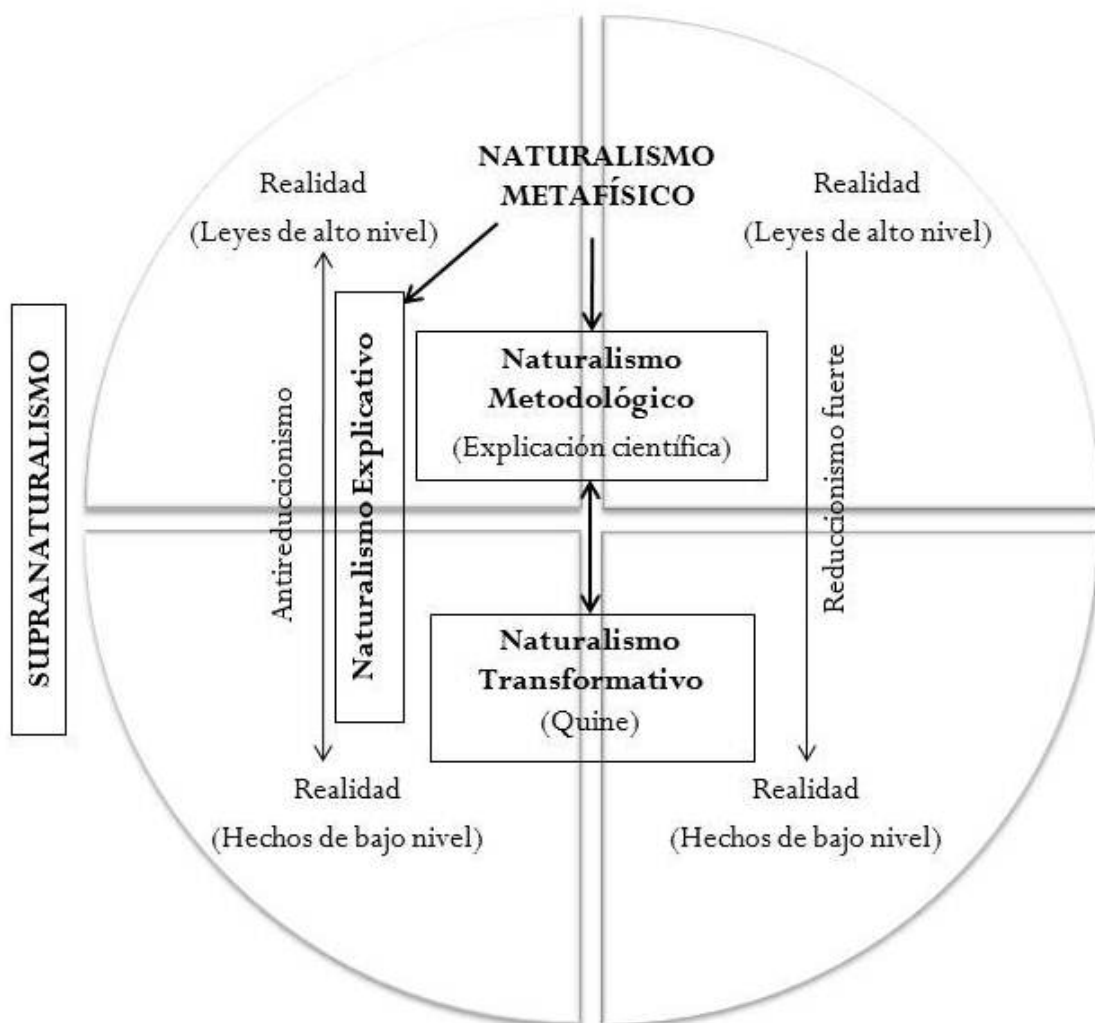


Figura 1.

Por otra parte, puesto que nuestro interés en el naturalismo obedece a la idea de defender una construcción naturalista de la moral, no podemos dejar de mencionar aquí las implicaciones que cada una de estas clases tiene para la normatividad. La conjunción de estas implicaciones constituirá, para Prinz, aquello que podemos denominar un naturalismo moral, tal y como se puede apreciar en la siguiente tabla (v. Tabla 2):

NATURALISMO MORAL	CLASES DE NATURALISMO	IMPLICACIONES PARA LA NORMATIVIDAD
	1. Naturalismo Metafísico	Supone que las normas morales, si existen, no requieren postular ninguna cosa que vaya más allá de lo que las ciencias naturales postulan (cf. Prinz, 2007a, 3).
	2. Naturalismo explicativo o descriptivo	Supone que podemos, en última instancia, describir cómo cualquier norma moral es realizada mediante entidades naturales (cf. ibíd.).
	3. Naturalismo metodológico	Supone que podríamos investigar las normas usando todas las herramientas provenientes de recursos empíricos (cf. ibíd.).
	4. Naturalismo transformativo	Supone que debemos investigar las normas desde adentro de nuestros sistemas actuales de creencias y, como resultado, las normas que actualmente aceptamos influirán en nuestras intuiciones acerca de cuáles normas podríamos sostener. Si elegimos cambiar nuestras normas, no podemos hacerlo adoptando una instancia trascendental que se agrupe arbitrariamente a las normas que actualmente aceptamos (cf. ibíd.).

Tabla 2.

### 1.1.1 El naturalismo moral frente a la Ley de Hume

¿Puede, entonces, el naturalismo moral impugnar la distinción es/debe o *Ley de Hume*?

Prinz piensa que no solo puede hacerlo, como veremos a continuación, sino que también la debe preservar. Pero antes veamos lo que dice Hume acerca de esta distinción en su *Tratado sobre la naturaleza humana*:

No puedo pasar por alto añadir a estos razonamientos una observación que, quizá, sea de importancia. En todo sistema de moral que hasta ahora he examinado he advertido siempre que el autor procede durante un lapso de tiempo según la manera usual de razonar, probando la existencia de Dios o haciendo observaciones sobre las cosas humanas; pero de repente me sorprende hallar que en vez de las cópulas ordinarias de las proposiciones *-es* y *no es-* me encuentro con que no aparece proposición que no esté conexas con un *debe* o un *no debe*. Este cambio es imperceptible, mas no obstante es de suma importancia hasta el final. Al expresar este *debe* o *no debe* algún tipo nuevo de relación o afirmación, es preciso que se observe y explique, a la par que se dé alguna razón de lo que parece del todo inconcebible, a saber, cómo esta nueva relación puede ser deducida de otras que son por entero diferentes de ella. Pero como de ordinario los autores no hacen uso de esta precaución, me permito advertírsele a los lectores. Estoy seguro que si se parara mientes a este punto nimio, los sistemas de moral

corrientes sufrirían subversión, y veríamos que la diferencia entre vicio y virtud no está fundada exclusivamente en relaciones de objetos ni se percibe por la razón (Hume, 1981, libro III, parte ii, sección 1).

Con esta distinción —Ley de Hume, en adelante— Hume establece la ilegitimidad de pasar de la descripción de un hecho a la formulación de un principio moral —esto es, que no es legítimo derivar un «debe» imperativo de un «es» indicativo—. A esto han recurrido numerosos críticos de la ética naturalista para argumentar que los planteamientos básicos de dicha ética están viciados de raíz. Sin embargo, Prinz articula una explicación acerca de la manera en que su proyecto naturalista impugna la Ley de Hume, al tiempo en que también considera que es posible preservar dicha ley, haciendo fuerte la idea —defendida por el mismo Hume— de que juzgar algo como moralmente malo es el resultado de “experimentar desaprobación” y juzgar algo como bueno es el resultado de “experimentar aprobación”.

#### *1.1.1.1 Impugnación de la Ley de Hume*

Para este propósito, Prinz ofrece lo que él denomina un argumento rápido y no concluyente sobre cómo se deriva un debe de un es, puesto que, en su opinión, una defensa completa del argumento requeriría una laboriosa excursión por la filosofía del lenguaje. La meta del autor es pues una más modesta, y consiste en “[...] indicar una manera en la cual un naturalista podría considerar hechos tanto morales como naturales (sostenidos por hechos descriptivos), pero también no irreductibles (y, así, ya no tan sostenidos)” (Prinz, 2007a, 4).

El primer paso consiste en señalar qué es un deber. Para esto, Prinz restringe la palabra “deber” al ámbito de la moralidad e intenta determinar el uso del concepto que expresa esta palabra, esto es, intenta determinar lo que la gente tiene en mente cuando dice que algo es obligatorio. El resultado, en palabras de Prinz, es el siguiente:

Desde la teoría que defiende, cuando una persona dice que un curso de acción es obligatorio, este juicio expresa lo que podría llamarse un sentimiento prescriptivo. Un sentimiento prescriptivo es una disposición emocional compleja: si uno tiene este sentimiento acerca de una forma de conducta particular, entonces uno está dispuesto a adherirse a esta conducta, y uno está dispuesto a sentirse mal si no lo hace; uno también está dispuesto a condenar a quienes no se adhieran a esta conducta (ibíd.).

Así, por ejemplo, si Smith juzga honestamente que uno podría dar por caridad, Smith expresa un sentimiento que lo dispone a sentirse mal si no da por caridad y se disgusta si otra persona no da por caridad. Y, en consecuencia, el significado de la palabra “deber” es un sentimiento prescriptivo, esto es, aquello que expresa la palabra. En una palabra, el significado de deber es el conjunto de condiciones que satisfacen el juicio de que algo es obligatorio. Una vez determinadas estas condiciones, Prinz esquematiza su argumento de la siguiente manera:

1. Smith tiene una obligación de dar por caridad si “Smith debe dar por caridad” es verdadero.
2. “Smith debe dar por caridad” es verdadero, si la palabra “deber” expresa un concepto que se aplica a la relación entre Smith y dar por caridad.
3. La palabra “deber” expresa un sentimiento prescriptivo.
4. Smith tiene un sentimiento prescriptivo frente a ‘dar por caridad’.
5. Así, la oración “Smith debe dar por caridad” es verdadera.
6. Así, Smith tiene una obligación de dar por caridad (Prinz, 2007a, 5).

De acuerdo con esto, Prinz defiende que a) la conclusión del argumento es un hecho prescriptivo; b) las premisas son descriptivas; porque c) la palabra “deber” ha sido mencionada, mas no usada; y, en consecuencia, d) la Ley de Hume ha sido impugnada. Aquí, cabe anotar que Prinz considera que está ofreciendo una teoría sustantiva, y no meramente convencional, del significado de términos normativos. Esto en razón de que, tal y como está expuesto el argumento, estamos evitando el riesgo de considerar a una persona como teniendo una obligación aun cuando ella no esté actualmente obligada —cuestión que no suele sortear el convencionalismo—.

Ahora bien, Prinz observa que su premisa 3, a diferencia de las demás, es muy controversial, y es por esto que señala que gran parte de su libro está encaminado a proveer argumentos que la hagan más convincente. Sobre este tema volveremos en el siguiente ítem. Por lo pronto, Prinz concluye que:

[...] el argumento puede ser modificado para acomodar otras teorías. Si el naturalismo es verdadero entonces los conceptos morales son vacíos o expresan propiedades que en última instancia pueden ser descritas sin vocabulario moral. Si mi análisis del deber es incorrecto, sustituyamos esto por otro análisis y reemplacemos la premisa 3 con la correspondiente descripción de los hechos morales subyacentes a la obligación. Ahora, revisemos la premisa 4 y el argumento correrá. Si hay obligaciones, entonces ellas pueden ser derivadas de esta forma puramente descriptiva sobre cualquier explicación naturalista (ibíd.).

#### *1.1.1.2 Defensa de la Ley de Hume*

Con todo, parece que el argumento de Prinz aún no hace justicia a la fuerte intuición que favorece la Ley de Hume, pues sigue pareciendo un error categorial pasar de premisas sobre cómo son las cosas a conclusiones sobre cómo deberían ser. Además, también parece que de una postura normativa como la que se sustenta en la premisa 3 podrían derivarse obligaciones inadmisibles. Así, por ejemplo, un asesino podría tener un sentimiento prescriptivo al asesinar personas y, de acuerdo con el argumento, parece sostenible que el asesino está obligado a matar. Prinz es consciente de estas dos preocupaciones, pero piensa que estas pueden ser matizadas sin cambiar diametralmente su argumento.

A la primera preocupación, Prinz responde reiterando que, efectivamente, la conclusión (6) del argumento ha derivado un hecho prescriptivo de premisas descriptivas (1-5) y que, de esta manera, parece indudable que se ha infringido la Ley de Hume. Pero, en otro sentido, él afirma que la Ley de Hume no ha sido quebrantada. Esto en razón de que la frase ‘hecho prescriptivo’ es ambigua, puesto que, según Prinz, “Sobre una lectura, un hecho prescriptivo es justamente un hecho sobre lo que alguien

está obligado a hacer. Pero, un hecho prescriptivo también puede ser interpretado como un juicio prescriptivo o, más sucintamente, como una prescripción” (Prinz, 2007a, 6). De acuerdo con esto, podemos y debemos interpretar la conclusión en el primer sentido, porque la conclusión dice que “Smith tiene la obligación de dar por caridad”, y no que “Smith debería dar por caridad”. Así, para Prinz, un “debe” expresa un sentimiento prescriptivo que solo puede ser usado por un hablante que tenga dicho sentimiento; pero en el argumento no hay ninguna premisa que implique que el orador tenga una disposición a reaccionar emocionalmente por caridad. Por lo tanto, de las premisas no se puede inferir un “debe” sino una “obligación”, que no necesariamente implica una prescripción, sino un hecho prescriptivo. La Ley de Hume, en consecuencia, aún está salvaguardada.

La segunda preocupación —sobre las posibles obligaciones inadmisibles que podrían derivarse de la premisa 3— se sostendría en que el argumento de Prinz viola la Ley de Hume, porque violando, a su vez, un principio semántico básico en él se pasa de la premisa semántica “Smith debería dar por caridad” es verdadero, a la exigencia de que Smith tiene una obligación de dar por caridad. Para hacer frente a esta objeción, el autor introduce la distinción entre una implicación semántica y una implicación conversacional —‘conversacional’ en cuanto a que la interpretación del juicio está sujeta al contexto mismo de una conversación—. Así, según Prinz, “estar obligado” implica “debería”, pero de una manera conversacional, mas no en un sentido semántico. Hecha esta distinción, Prinz señala que la objeción descansa en el supuesto de que si “Smith debería dar por caridad es verdadero”, entonces Smith debería dar por caridad —que sería un ejemplo del principio de des-entrecorillado—. Sin embargo, él llama la atención sobre el hecho de que no siempre podemos inferir P de “P es verdadero”, puesto que el des-entrecorillado no siempre es permitido cuando usamos indexicales

tales como “yo” —así, por ejemplo, si suponemos que Smith pronuncia la oración “yo soy Smith”, la oración será verdadera, pero de aquí no se sigue que yo soy Smith sea verdadero— (cf. Prinz, 2007a, 6-7).

El argumento de Prinz es pues un contraejemplo al principio de des-entrecorillado. Para él, “[...] “deber” es como un indexical en el que su significado no es agotado por su contribución a una proposición expresada” (Prinz, 2007a, 7), esto es, el significado de “deber” requiere de interpretaciones que no se agotan en el carácter semántico de la palabra, sino que, dicho sea de paso, son conversacionales. Prinz espera que así como el caso de “yo” muestra que el des-entrecorillado tiene excepciones bien conocidas, también “deber” puede ser una excepción, lo cual hace que el argumento se sostenga.

Pero tratemos de entender mejor a qué se refiere Prinz con el hecho de que “debe” es una implicación conversacional de “obligación”, mas no una implicación semántica. Para él, nuestra incomodidad al afirmar que las personas tienen obligaciones que no aprobamos —ya que nadie quisiera aprobar que, por ejemplo, los asesinos estén obligados a matar— tiene un origen pragmático: “Las atribuciones de obligaciones conversacionalmente implican juicios prescriptivos. Si yo digo que alguien está obligado a dar por caridad, probablemente tengo un interés en transmitir mi forma de sentir” (ibíd.). En otras palabras, afirmar la existencia de una obligación es un modo de transmitir que yo pienso que la persona debería hacer algo; pero es muy diferente decir “Smith debería dar por caridad, aunque él no debería dar por caridad” a decir “Smith tiene una obligación de dar por caridad, pero él no debería”. Las dos frases parecen contradictorias, pero la segunda, a diferencia de la primera, se puede defender con coherencia, puesto que conversacionalmente podríamos admitir que “[...] los soldados japoneses de la Segunda Guerra Mundial tenían una obligación de sacrificar sus vidas



como pilotos kamikazes, pero ellos no deberían haber hecho esto” (ibíd.). De aquí que Prinz concluya que:

Las obligaciones pueden ser deducidas de premisas descriptivas, pero ellas no tienen que ser endosadas por quienes las deducen. Los endosos simplemente son implicados [de manera conversacional]. Ellos no pueden ser deducidos [de manera semántica]. Creer que Smith debería dar por caridad requiere hacer un juicio prescriptivo. Para hacer una prescripción, necesitamos estar en un estado psicológico particular —necesitamos prescribir—. Este es el sentido en el cual no podemos derivar un “deber” de un “es” (ibíd.)<sup>1</sup>.

#### 1.1.1.1 *Defendiendo el subjetivismo*

Esta conclusión, y ante todo la afirmación de que prescribir es estar en un estado psicológico particular, le permite a Prinz afirmar que la moralidad es subjetiva. Para él, si “deber” expresa un sentimiento prescriptivo, que es interpretado conversacionalmente, entonces conceptos morales tales como “deber” son conceptos fundamentalmente subjetivos —y lo mismo ocurre con conceptos morales tales como “bueno” y “malo” o “correcto” e “incorrecto”—.

La idea es pues derivar prescripciones de descripciones, esto es, derivar hechos metafísicos de hechos psicológicos, puesto que, en esta interpretación, correcto e incorrecto serán los referentes de nuestros conceptos de “correcto” e “incorrecto”. Es así como Prinz espera que si el análisis de nuestros conceptos morales descubre una conexión fuerte con respuestas subjetivas, entonces estos conceptos pueden hacer referencia a algo subjetivo. Para él: “La psicología moral implica hechos acerca de una ontología moral y una psicología sentimental puede implicar una ontología subjetiva” (Prinz, 2007a, 8).

---

<sup>1</sup> Las aclaraciones entre corchetes son mías.

### 1.1.2 El naturalismo moral y la ‘falacia naturalista’

Una postura muy similar a la Ley de Hume, y que también se dirige en contra del naturalismo moral, fue la que defendió G. E. Moore en su *Principia Ethica* (cf. Moore, 1997). Él trató de refutar la tesis del naturalismo ético mediante la formulación de dos argumentos: la defensa del carácter simple de la bondad y el famoso argumento de la ‘pregunta abierta’. Con el primero, que intenta llegar a una conclusión de tipo metafísico, Moore pretendía defender que la propiedad de ser bueno no está constituida por ninguna otra propiedad; con el segundo, un argumento de tipo semántico, él rechazó cualquier propuesta de definición de ‘bueno’. El segundo argumento, como espero mostrar a continuación, descansa sobre un error, y el primero, por su parte y en tanto que se apoya en el segundo, también incurre en un error, como es el de pretender derivar conclusiones metafísicas de premisas de tipo semántico.

#### 1.1.2.1 El carácter simple de la bondad

Para conservar el orden de la argumentación de Moore, me parece adecuado comenzar con su argumento a favor del carácter simple e indefinible de la bondad. A este respecto, en su *Principia Ethica*, Moore afirma:

Ahora, sin entrar en la discusión del sentido propio de la palabra [...], puedo decir que me propongo usar ‘ética’ [...] para abarcar una investigación que no cuenta, en todos los casos, con otra palabra: la investigación general de qué sea bueno [...] Y al examen de esta pregunta le doy el nombre de ética, puesto que tal ciencia debe incluirla siempre (Moore, 1997, §2).

De acuerdo con esto, Moore parte del supuesto de que el propósito principal del estudio ético debe ser aquello que es común a todos los juicios éticos, es decir, lo que los hace pertenecer al ámbito de la ética. Esta empresa, en opinión de Moore, requiere determinar a qué cosas aplicamos usualmente el adjetivo ‘bueno’, pero no

podemos indagar este tipo de cosas, aunque ellas sean objeto de la ética, sin antes responder qué significa ‘bueno’ o sin que contemos con una definición de ‘bueno’:

Es una inquisición a la que debe dirigirse la más especial atención; puesto que la interrogación acerca de cómo definir ‘bueno’ es la más fundamental de toda la ética. Lo que se entiende por ‘bueno’ es, de hecho, el único objeto simple del pensamiento que es peculiar de la ética [...]. A menos que esta primera pregunta se entienda plenamente y se reconozca su respuesta correcta, de modo claro, el resto de la ética será inútil (Moore, 1997, §5).

Para responder a esta pregunta, que Moore encuentra fundamental en la investigación ética, él señala entonces que hay tres tipos de definición: una ‘verbal arbitraria’, una ‘propriadamente verbal’ y una que señala la ‘constitución’ o naturaleza del objeto. La primera definición, de carácter estipulativo, apunta a lo que cada persona entiende cuando dice que algo es ‘bueno’. La segunda, de carácter convencional, se refiere al uso habitual que la gente hace del adjetivo ‘bueno’, en consonancia con las reglas aceptadas para el uso de este término. La tercera, por su parte, está encaminada a establecer una definición real de ‘bueno’, que señale la verdadera naturaleza del término, y es esta la definición en la que Moore está interesado. Sin embargo, él señala que:

Puede darse una definición de caballo, porque un caballo tiene muchas propiedades y cualidades diferentes que pueden ser enumeradas. Pero cuando se han enumerado todas, cuando se ha reducido un caballo a sus términos más simples, luego estos términos ya no pueden definirse. Son simplemente algo que se puede pensar o percibir y a quien no puede pensarlos o percibirlos, no es posible hacerle conocer su naturaleza por medio de ninguna definición” (Moore, 1997, §7).

De esta manera, la definición de ‘bueno’ que Moore busca es una definición analítica, en tanto que se trata de dar la enumeración de las partes de la cosa que debe ser definida y las relaciones existentes entre ellas. Pero esta última definición resulta imposible para un objeto simple y, en opinión de Moore, ‘bueno’, como común denominador de todos los juicios éticos, es un concepto simple. En consecuencia, Moore llega a la conclusión de que la naturaleza de ‘bueno’ es indefinible. La postura

de Moore es pues atomista, ya que supone que ningún objeto simple puede ser analizado o definido; y, por otro lado, se trata también de una postura intuicionista, en tanto que esta imposibilidad de definir la palabra “bueno”, conlleva que únicamente podamos intuirlo. De acuerdo con esta postura las propiedades éticas, como la bondad, no son propiedades naturales (es decir, no son ubicables en el tiempo ni en el espacio) y, por ende, no se pueden conocer bajo las directrices del naturalismo moral que expusimos anteriormente, sino por una facultad cognoscitiva especial: la intuición.

De este modo, para Moore, si ‘bueno’ es pues una propiedad que solo puede intuirse, no debe confundirse con otras propiedades de tipo natural, tales como ‘útil’, por ejemplo. Las cosas buenas también podrían ser útiles, pero esto no significa que la propiedad de ser bueno equivalga a la de ser útil. Según él, este error de confundir dos propiedades es común, muchos filósofos “han pensado que, cuando nombran esas otras propiedades, están definiendo ‘bueno’ realmente, y que no son, de hecho, ‘otras’ sino absoluta y enteramente iguales a la bondad” (Moore, 1997, §10). Al hecho de incurrir en este error es a lo que Moore se refiere como ‘falacia naturalista’. Para él, las cosas buenas pueden ser algo más, pero es un error pensar que esta otra propiedad o propiedades que acompañan a lo bueno sean lo mismo que la bondad. Esta identificación entre la bondad y otras propiedades no es legítima, pues una cosa es la co-extensividad de dos propiedades, y otra muy diferente la identidad de las mismas: uno puede descubrir que dos propiedades van siempre juntas, mas esto no significa que sean la misma. Así, dicho en términos de Moore:

Quando alguien confunde entre sí dos objetos naturales, definiendo el uno en lugar del otro, si, por ejemplo, se confunde a sí mismo -un objeto natural- con ‘place’ o ‘placer’ -que son otros objetos naturales-, no hay entonces razón para denominar a esto falacia naturalista. Pero si confunde ‘bueno’, que no es, en el mismo sentido, un objeto natural, con cualquier objeto natural, hay razón entonces para llamar a esto falacia naturalista; el que se dé con relación a ‘bueno’ la señala como algo muy específico, y este error específico requiere un nombre por

ser tan habitual (...). Pero, por ahora, basta señalar que aún si fuera un objeto natural, esto no alteraría la naturaleza de la falacia ni disminuiría un ápice su importancia (Moore, 1997, §12).

En suma, la falacia naturalista es pues la conjunción entre: *i.* el intento por definir la bondad; y *ii.* el intento por identificar la bondad con otras propiedades de carácter natural que están asociadas con ella.

#### 1.1.2.2 *El argumento de la pregunta abierta*

Moore ha argumentado a favor del carácter simple de la bondad y de que esta solo es cognoscible mediante la intuición; sin embargo, él considera que es necesario dar argumentos adicionales que apoyen su postura, ante todo en lo que respecta a la idea de que lo bueno es una propiedad indefinible. Para Moore, si ‘bueno’ no denota algo simple, entonces es preciso probar que ‘bueno’ denota algo complejo. Pero para probar que esto último no es el caso, él formuló el ya muy conocido argumento de ‘la pregunta abierta’.

Este argumento señala que para cualquier definición que se plantee de ‘bueno’, siempre tendrá sentido preguntarse si tal definición es correcta, es decir, siempre podremos preguntarnos si lo que se ofrece como definición de ‘bueno’ es bueno realmente y entender tal pregunta como una pregunta que tiene sentido hacer y de la que, por tanto, se dice que está abierta. Así, por ejemplo, si se define ‘bueno’ como ‘placentero’, y vemos luego que tiene sentido preguntarse si lo placentero es bueno, entonces esto indica que ‘bueno’ y ‘placentero’ no son sinónimos. En este sentido, ‘placentero’ no puede significar lo mismo que ‘bueno’, ya que preguntar ‘¿lo placentero, es placentero?’ no es lo mismo que preguntar ‘¿lo placentero es bueno?’. La primera pregunta constituye una tautología y, en consecuencia, es una pregunta cerrada, cuya respuesta es lógicamente *sí*. La segunda pregunta, en cambio, es una

pregunta abierta, no tautológica y con sentido, pues es lógicamente posible que no todas las cosas que producen placer sean buenas.

Moore concluye así que ‘bueno’, en tanto que no tiene sinónimos, no puede ser definido en modo alguno, puesto que cualquier definición incurrirá en una pregunta abierta. Así, por ejemplo, bajo una postura hedonista nadie aceptaría que ‘lo que produce placer es bueno’ es una tautología igual a ‘lo que produce placer produce placer’, porque ‘bueno’ y ‘placentero’ no significan lo mismo; si se tratara de una relación de sinonimia, la pregunta en cuestión debería ser cerrada. En consecuencia, la propiedad de ser bueno es una propiedad simple, puesto que, en contraste con las propiedades complejas, no está sujeta a definición.

Así pues, Moore defiende que la bondad es un objeto simple e indefinible y que se trata de una propiedad no-natural, solo cognoscible mediante la intuición. De aquí que él afirme que sería un error intentar definir la bondad en términos naturales, esto es, incurrir en la ‘falacia naturalista’. Además, él formula su argumento de la pregunta abierta, para demostrar, por reducción al absurdo, que ‘bueno’ no tiene sinónimos. De esta manera, Moore consolida su tesis ontológica de la simplicidad de la bondad y refuerza sus críticas al naturalismo moral.

#### *1.1.2.2.1 El error categorial en ‘la pregunta abierta’*

Ahora bien, como lo sugerí anteriormente, el argumento de Moore descansa sobre un error, puesto que a partir de argumentos conceptuales está arguyendo a favor de una conclusión metafísica: él parte del hecho de que es una pregunta abierta preguntarse si cualquier propiedad P es buena, como soporte para llegar a la conclusión de que la propiedad de ser bueno no está constituida por ninguna otra propiedad. Esto es

indudablemente un error, un error que podríamos denominar ‘categorial’, puesto que de manera arbitraria pasa del plano conceptual al plano ontológico.

Para aclarar mejor en qué consiste este error, pasemos ahora a examinar la crítica de Prinz a los argumentos de Moore. Prinz considera que nosotros podemos reconocer la existencia de identidades de manera *a posteriori* —hechos metafísicos que no pueden ser descubiertos mediante el análisis conceptual—. Así, por ejemplo, antes de la química moderna nadie sabía que el alcohol era un compuesto de hidróxido, con una estructura molecular  $C_nH_{2n+1}OH$ , pero la afirmación de que el alcohol es un compuesto de hidróxido no es ciertamente una verdad obvia, porque uno puede tener el concepto de alcohol sin saber nada de química. Sin embargo, afirma Prinz, así como sería un error desafiar esta identidad señalando que uno puede preguntarse inteligiblemente si esto es verdad, fue un error de Moore inferir que la bondad es una propiedad no natural del hecho de que siempre cabrá preguntarse inteligiblemente si esto es así (cf. Prinz, 2007a, 39).

De esta manera, el punto de Prinz está en que “los argumentos de pregunta-abierta no pueden establecer conclusiones metafísicas, aunque puedan establecer conclusiones conceptuales” (ibíd.). Esto en razón de que si es una pregunta abierta preguntarse si alguna propiedad natural  $N$  es buena, entonces el concepto de  $N$  no puede ser parte del concepto de lo bueno. Esta es una conclusión conceptual. Así, por ejemplo, si es una pregunta abierta preguntarse si la propiedad de lo útil es buena, entonces el concepto de lo útil no puede ser parte del concepto de lo bueno, porque no hay una relación conceptual entre los dos. Sin embargo, en la práctica, el concepto de la bondad puede referirse a la propiedad de la utilidad sin que haya un enlace conceptual entre los dos, y entender el concepto de lo bueno puede implicar entender la propiedad de la utilidad.

Para explicar esto tomemos un caso mucho más claro. Es parte del concepto corriente de alcohol (es decir, el concepto usado por la mayoría de los adultos cuando con frecuencia piensan en el alcohol) que beberlo puede acarrear una intoxicación. Si esto es parte del concepto corriente, entonces nadie que comprenda ese concepto podría inteligiblemente preguntarse si beber alcohol puede acarrear intoxicación. Aquí hay un claro enlace conceptual y, por ende, no hay ninguna pregunta abierta sino una cerrada. En el caso moral, si fuera parte del concepto corriente de bueno (esto es, el concepto usado por la mayoría de las personas cuando piensan en lo bueno) propiciar utilidad, entonces nadie que comprenda ese concepto podría inteligiblemente preguntarse si lo bueno propicia la utilidad. Es discutible si es este el caso o no, y lo más probable es que no, pero el punto está en que una cosa es la conclusión conceptual y otra muy diferente la conclusión ontológica.

Para Prinz, el argumento de Moore también puede ser comparado y contrastado con el célebre argumento de Frank Jackson en contra del materialismo.

Jackson nos hace imaginar a una brillante neurocientífica llamada María, que es encerrada en una habitación negra y blanca. María sabe todo sobre lo que ocurre en el cerebro cuando las personas ven el rojo, pero esto no le ayuda a entender qué sería semejante a tener una experiencia del rojo. De la misma manera que Moore, Jackson trata de obtener una conclusión metafísica. Él dice que la experiencia de ver el rojo no puede ser idéntica a un estado cerebral. Esa conclusión es muy controversial. Muchos comentaristas piensan que es falaz inferir una conclusión metafísica acerca de la base de la experiencia del rojo a partir de premisas epistemológicas o semánticas acerca de si podemos inferir que el rojo es semejante al conocimiento del cerebro. Por eso no podemos admitir que el argumento de Jackson establezca una conclusión epistemológica. Los conceptos involucrados en saber qué experiencias son semejantes difieren de los conceptos involucrados en saber acerca de los estados cerebrales como tal: uno no puede inferir una clase de conceptos de las otras (Prinz, 2007a, 40).

De acuerdo con esto, los conceptos invocados en las teorías éticas normativas son diferentes de los conceptos de correcto e incorrecto. Uno no puede inferir que una acción es moralmente correcta o incorrecta de una descripción kantiana o milliana de



esta acción. Prinz encuentra que esto es suficiente para establecer que Kant, Mill y otros éticos normativos dejan de explicar los conceptos morales. Aunque, por supuesto, esto no prueba tampoco que el naturalismo moral metafísico sea correcto, pues no podemos inferir que Kant y Mill estuvieron equivocados sobre las propiedades morales del hecho de que ellos estuvieron equivocados sobre los conceptos morales. El punto está en que es posible, sin embargo, que los conceptos morales designen propiedades kantianas o millianas, pero concebir estas propiedades no es suficiente para captar conceptos morales. En una palabra, uno puede comprender qué es maximizar una utilidad y qué es universalizar una máxima conductual, al tiempo en que no comprende lo que las personas normalmente captan cuando usan las palabras “correcto” o “incorrecto” (cf. *ibíd.*).

De acuerdo con esto, Prinz encuentra que fue un error por parte de Moore pensar que ningún análisis de los conceptos morales permitiría una pregunta cerrada, y llegar a la conclusión de que los conceptos morales eran conceptualmente primitivos. Para Prinz, es perfectamente plausible suponer que alguien desarrolla la capacidad para la empatía, que empieza a sentirse espontáneamente feliz cuando otros están felices y triste cuando otros están tristes, que obtiene un deleite especial maximizando lo bueno porque sabe que traerá la felicidad a otros, que se siente muy culpable cuando no maximiza lo bueno porque sabe que dejó de brindar placer a alguien a quien podría haber hecho más feliz y que experimenta enfado hacia otras personas cuando dejan de maximizar la utilidad. En este punto, de acuerdo con Prinz, parece igualmente plausible decir que esta persona cree que es incorrecto no maximizar la utilidad y correcto maximizar la utilidad, es decir, podemos atribuirle creencia moral, una vez empieza a tener una actitud emocional hacia la maximización de la utilidad (cf. Prinz, 2007a, 41 ).

Esto significa que nosotros atribuimos actitudes morales a una persona en cuanto tenga actitudes emocionales, puesto que una vez que ciertas acciones le permiten sentirse culpable e indignada, decimos que ella tiene un punto de vista moral. Y es en este sentido que Prinz concluye que hay un enlace conceptual entre los juicios morales y las respuestas emocionales, puesto que atribuimos actitudes morales sí y sólo si una persona tiene determinadas respuestas emocionales. Para Prinz:

Esta tesis positiva puede ser expresada diciendo que, en cuanto una persona adquiere actitudes emocionales, ciertas preguntas se vuelven cerradas. Pero uno tiene que ser cuidadoso al decir qué es una pregunta cerrada. Suponga que alguien se sintiera culpable si robara e indignada si otra persona robara. Aquí no parece haber una pregunta acerca de si esta persona tiene una actitud moral hacia robar. Claramente ella mira el robo como incorrecto. La pregunta por la actitud de esta persona es cerrada. Si tengo razón, Moore estaba equivocado cuando afirmó que los conceptos morales no admitían análisis. Los conceptos morales corrientes están conceptualmente vinculados con las emociones, y esto nos permite decir que alguien que alberga culpa y enfado está tomando una postura moral (ibíd.).

Ahora bien, ante esta tesis, Prinz considera la posible objeción de que el enlace entre los conceptos morales y las emociones nos deja con preguntas abiertas. Esto en razón de que, aun suponiendo que alguien siente indignación hacia aquellos que roban, uno podría preguntarse si robar es *realmente* incorrecto, esto es, preguntarse si la actitud es justificada. Sin embargo, a la pregunta de si esta pregunta abierta mina la afirmación de que hay un enlace conceptual entre lo incorrecto y ciertas emociones, Prinz responde con un rotundo no. En su opinión, la razón de por qué esta persona puede preguntarse si robar es incorrecto se refiere más a una brecha entre sentido y referencia, puesto que muchos de nuestros conceptos son comprendidos por significaciones de caracteres que no son ni necesarios ni suficientes para su admisión en una categoría. Así, por ejemplo, concebimos a las aves como criaturas con alas y plumas, pero podría haber un ave desalado e implume. Por consiguiente, con frecuencia podemos preguntarnos si un concepto que estamos aplicando, en cualquier ocasión en particular, es *realmente* aplicable. Así, en nuestro caso particular, la persona

en cuestión sabe que ella *concibe* el robo como malo, pero puede preguntarse si robar realmente cae bajo su concepto de malo. En tanto que sentido y referencia vienen aparte, a menudo hay preguntas abiertas sobre a qué se refiere un concepto dado, incluso si sabemos que el sentido por el que comprendemos el concepto ha sido satisfecho. De acuerdo con esto, el punto de Prinz está en que la pregunta abierta en este caso es una pregunta acerca de a qué se refieren los conceptos morales, no sobre su sentido, pero el hecho de que podemos decir que esta persona está moralizando sólo en virtud de sus actitudes emocionales indica que hay un enlace conceptual entre los conceptos morales y las emociones (cf. *ibíd.*).

## 1.2 EL ENFOQUE NATURALISTA

Hasta aquí, tomando como ejemplo particular el ámbito de la moral, he intentado sistematizar los argumentos de Prinz a favor de la defensa de un enfoque naturalista-cognitivo de los juicios de valor, frente a sus principales contraargumentos. Esto con el propósito de dar una base teórica legítima a este tipo de enfoque y de sugerir la viabilidad que las propuestas no naturalistas de la ética podrían encontrar en él. En esta convergencia, los límites rígidos e insoslayables entre el trabajo filosófico y el científico se vienen abajo, y aflora así la necesidad del trabajo interdisciplinario; porque no es lícito que ningún tipo de información y enfoque metodológico puedan ser deslegitimados de manera *a priori*.

Así, si descubrimos que como consecuencia de un determinado tipo de daño cerebral se ha perdido o deteriorado una determinada habilidad relacionada con los juicios de valor, mientras que otra aparentemente similar permanece intacta, nos vemos obligados a buscar una diferencia significativa correspondiente a las dos habilidades. Estos hallazgos vienen del trabajo psicológico empírico y la filosofía los

complementa mediante su interpretación e incorporación en una teoría general. De esta manera, bien harían los teóricos tradicionales en ética y estética en incorporar las evidencias de la psicología empírica a sus reflexiones, tal y como lo hacen, valga decirlo, muchos de los psicólogos con el trabajo filosófico. Ninguna de las dos áreas pierde su estatus con esta forma de trabajo interdisciplinario y, por el contrario, se fortalecen cada vez más en su objetivo común del progreso cognoscitivo. La ciencia, la reflexión filosófica, la práctica, la percepción, y las diferentes artes y moralidades no van por caminos diferentes, sino por una única senda que conduce al conocimiento y a formarse una idea del mundo circundante. Nuestro precepto fundamental debe ser el de perseguir el entendimiento de este mundo circundante rechazando la idea de que hay cosas que no podemos entender. Esto es lo que hemos denominado aquí un naturalismo metodológico, de acuerdo con el cual las explicaciones acerca de los hechos naturales se pueden hacer con medios igualmente naturales. Esta postura rechaza, de entrada, que en los ámbitos de la moral y de la estética debamos aceptar explicaciones de tipo supra-natural, porque estos ámbitos se derivan de nosotros y están, en consecuencia, anclados en la naturaleza humana y no en un reino espiritual. Los conceptos mismos son entidades naturales y ellos pueden ser investigados usando procesos naturales.

Estoy de acuerdo con Prinz en que el análisis conceptual es, como todas las herramientas legítimas de investigación, un método empírico, pero que, como tal, no resulta especialmente poderoso. El análisis conceptual procede mediante el acceso en primera-persona a estructuras psicológicas, o introspección. La introspección está propensa al error y es metodológicamente restringida, en tanto que se asocia con la obtención de conclusiones que usan un sujeto particular (uno mismo). El naturalismo, en contraste, sugiere investigar los conceptos usando herramientas de la ciencia social,

ya que si los conceptos son entidades naturales, entonces ellos surgen en formas naturales. Así, por ejemplo, los conceptos pueden ser adquiridos mediante la experiencia y ellos pueden ser revisados mediante la experiencia. Ellos no tienen un estatus especial cuando vienen a revelar hechos acerca del mundo.

Nuestra aproximación a los juicios morales y de gusto estético debe ser pues naturalista en todos los sentidos en que hemos caracterizado esta noción. Ante todo, se trata de un naturalismo metodológico que se sirve de los hallazgos empíricos provenientes de la neurociencia, la psicología, la antropología y la etología, entre otros. No se trata de rechazar tampoco métodos filosóficos no naturalistas, tales como el análisis conceptual, sino de complementarlos con la evidencia empírica, de tal manera que podamos movernos de los datos a la teoría mediante la sistematización de resultados en diferentes y coherentes corpus que pueden guiar una investigación futura.

A este respecto y como ejemplo de la fuente de dichos datos, permítaseme hacer una breve referencia a la neurociencia cognitiva social, que está emergiendo como un sub-campo de la neurociencia cognitiva (cf. Brothers, 2002). Su principio fundamental es que el ser humano usa mecanismos neuronales-cognitivos específicos para responder a las señales de su entorno. Se busca identificar así los componentes de la cognición social y relacionarlos con circuitos cerebrales específicos. Estudiando las habilidades sociales y los cerebros de primates no-humanos, se espera descubrir cómo se desarrolla la cognición social en los humanos. Por lo general, en esta tarea se asume que el cerebro tiene alguna potencia innata social y se indaga en la manera en que los cerebros jóvenes interactúan con el entorno durante el desarrollo para producir capacidades sociales adultas. La psicología del desarrollo, neurociencia clínica, anatomía comparativa y la primatología contribuyen en este trabajo. La dirección de la atención y ciertas configuraciones de los músculos faciales son señales usadas por muchos

primates. Un sistema neuronal evolutivamente viejo, presente en el cerebro humano normal en el nacimiento, puede preparar a un niño para responder a tales señales. Este sistema neuronal que traemos incorporado es el andamiaje para la acumulación de experiencias sociales subsiguientes, que son construidas sobre respuestas simples y anteriores a los gestos y expresiones de los rostros —a este respecto, como una herramienta práctica para familiarizarse con el lenguaje de la neurofisiología, en particular en cuanto a su relación con los procesos que determinan el juicio moral, sugiero ver el esquema general de Joshua Greene y Jonathan Haidt sobre las partes del cerebro involucradas en estos procesos; material incluido aquí como apéndice—.

---

## 2. LA TESIS EMOCIONALISTA EN LA EXPLICACIÓN COGNITIVA DEL JUICIO MORAL

La indagación acerca de la naturaleza de los juicios morales ha cobrado fuerza, cada vez más creciente, en los enfoques cognitivos de la filosofía moral contemporánea. El problema subyacente a esta indagación es el que parte de la cuestión acerca de la manera en que se definen los juicios y reglas morales o acerca de qué características los pueda distinguir de otro tipo de juicios y reglas. Sin embargo, siguiendo a Nado et al., tal parece que en los trabajos recientes de psicología empírica esta cuestión ha sido cambiada por la pregunta acerca de cuál es la base principal de los juicios morales (cf. Nado et al., 2009, 1). Esto en razón de que dichos trabajos sugieren que la moralidad es una clase natural, y la evidencia que arrojan parece comenzar a descubrir las que serían las propiedades esenciales de las reglas morales.

La filosofía ha vuelto así su mirada a la psicología empírica y, para no perder sus raíces, ha revivido con estos trabajos la muy conocida discusión de la tradición moderna entre emocionalistas y racionalistas. Hume defendió que la razón es “esclava de las pasiones” y que, en consecuencia, los juicios morales se derivan de las emociones

morales; mientras que Kant sostuvo que todos los requerimientos morales deben ser derivados de un principio racional (el imperativo categórico). De manera similar, hoy en día los filósofos cognitivos se preguntan si, a partir de la evidencia empírica, podemos afirmar que los juicios morales están constituidos principalmente sobre la base de la emoción, o son primariamente producto de la razón.

Esta pregunta, planteada así en términos de una disyunción excluyente, parece dividir al grupo de filósofos morales de corte cognitivista en dos bandos: humeanos y kantianos. En este debate, los primeros se apoyan en el cada vez más preciso mapeo de las partes del cerebro relacionadas con la emoción que se activan cuando emitimos juicios morales. Por su parte, el segundo grupo guía sus indagaciones mediante la aplicación de pruebas psicológicas, encaminadas a mostrar que un remanente de tipo racional se encuentra presente cada vez que emitimos juicios morales. De este modo, unos y otros interpretan los hallazgos empíricos de tal manera que apoyen sus enfoques, bien sea uno corporal, protológico o emocional, o bien sea un enfoque racional, lógico o conceptual.

En líneas muy generales, considero pertinente hacer aquí una presentación esquemática de los actores principales de este debate, antes de entrar a examinar sus posturas con mayor detalle. Así, en el primer grupo y como defensores principales del emocionalismo, se encuentran Jonathan Haidt (cf. Haidt, 2001) y Jesse J. Prinz (cf. Prinz, 2006 y 2007a). A la postura de Haidt sugiero denominarla “Emocionalismo débil”, en tanto que admite, en cierta medida, la existencia de casos en los que la razón es suficiente para emitir un juicio moral; a la postura de Prinz, por su parte, sugiero denominarla “Emocionalismo fuerte”, ya que defiende, propiamente hablando, la necesidad y suficiencia de las emociones para emitir juicios morales. Este tipo de emocionalismo es más radical que el anterior, en tanto que en él aquellos juicios que



podríamos denominar ‘desapasionados’ —porque parecieran utilizar como fuente primaria y suficiente a la razón— son realmente subsidiados por los juicios morales genuinos, esto es, por los emocionales. Por otra parte, el bando del racionalismo, o innatismo, está integrado por Susan Dwyer (cf. Dwyer, 1999) y Arnold Hauser (cf. Hauser, 2006), quienes defienden la existencia de una facultad gramática moral universal, que consistiría de los principios y parámetros, racionales e innatos, que serían indispensables para la producción del juicio moral. Y, finalmente, existe también una tercera postura en este debate que, a mi parecer, podría interpretarse como una conciliación de los dos grupos antes mencionados. Se trata de la postura de Joshua Greene (cf. Greene et al., 2001), según la cual las emociones cumplen un papel importante en el proceso de la emisión de los juicios morales y, sin embargo, el razonamiento también juega un papel nada despreciable en este proceso.

De acuerdo con esto, cabe preguntarse: ¿es posible que tanto la razón como la emoción cumplan un papel igualmente importante en la producción de los juicios morales? A esta pregunta, en mi opinión, no podemos responder con un sí o un no, de una manera contundente, pero sí podemos sugerir el sí señalando que, al descartar la posibilidad misma a la que apunta esta pregunta, el debate está viciado desde el inicio mismo de la discusión.

Dicho esto, la exposición de este capítulo se encamina a examinar si en efecto el debate entre emocionalistas y racionalistas —o innatistas— incurre en un radicalismo unilateral. Esto con el propósito de evaluar la solidez de los argumentos emocionalistas, que son presentados por Prinz como el argumento a la mejor explicación; y, dicho sea de paso, mi intención principal con este examen es socavar la pretensión de suficiencia del emocionalismo de Prinz, haciendo fuerte la idea de Greene de que tanto la razón como la emoción cumplen un papel igualmente importante en la producción de los

juicios morales. Siendo este último el propósito principal de este capítulo, debo aclarar que en la exposición del debate en general utilizaré la postura de Prinz como eje conductor. Espero mostrar, así, que en el proceso de emisión de los juicios morales todavía podemos defender una vieja intuición kantiana, según la cual la sensibilidad sin el entendimiento es ciega y el entendimiento sin la sensibilidad es vacío.

## 2.1 EL EMOCIONALISMO MORAL

El gran dominio que, en el siglo XX, tuvo la postura racionalista en el debate acerca de qué, entre la razón y la emoción, cumple un papel fundamental en la producción de los juicios morales, parece haber llegado a su fin con la psicología reciente. Esto en razón de que la postura emocionalista comenzó a fortalecerse con la radical y provocativa propuesta del modelo “intuicionista social”, del filósofo y psicólogo Jonathan Haidt (cf. Haidt, 2001). Sin embargo, pienso que ha sido Jesse J. Prinz el autor que ha desarrollado dicho enfoque de la manera más completa, coherente y fundamentada en la filosofía moral contemporánea (cf. Prinz, 2006 y 2007a). Por esta razón, me centraré aquí en la exposición de las tesis principales de estos dos autores, haciendo un especial énfasis en el segundo, en procura de ofrecer la visión más completa del emocionalismo moral e intentar, así, evaluar su poder explicativo.

### 2.1.1 El modelo “intuicionista social” de Jonathan Haidt

De acuerdo con el modelo “intuicionista social” de Haidt, las capacidades emocionales involucran afectos e intuiciones que se ocupan de llevar a cabo casi todo el proceso de la producción de los juicios morales. Para Haidt, el carácter ‘social’ de este modelo se debe a que desestima el razonamiento privado de las personas en la emisión del juicio moral y, en contraste, destaca la importancia de las influencias sociales y culturales en esta emisión; por otro lado, se trata de un modelo ‘intuicionista’, en tanto que afirma

que el juicio moral, por lo general, es el resultado de valoraciones rápidas y automáticas que se dan, así, a manera de intuiciones (cf. Haidt, 2001, 814). La razón, en este modelo, queda relegada entonces a hacer el papel de abogado o agente de relaciones públicas, cuyo fin principal es el de ofrecer justificaciones *post-hoc* de los juicios morales, una vez que estos ya se han hecho.

Ahora bien, en el modelo de Haidt, el primer paso para que surja un juicio moral es la percepción de un evento moralmente relevante. Después, son las intuiciones morales las que, de manera rápida y espontánea, surgen en una persona A y responden a la situación moral. Estas intuiciones no requieren de ningún tipo de razonamiento, sino que deben ser entendidas, en la propuesta de Haidt, como reacciones afectivas, tales como el enfado o disgusto, por ejemplo. De esta manera, en una situación normal en la cual el sujeto no requiera hacer ningún tipo de justificación, para otros (para una persona B, por ejemplo) o para sí mismo, el proceso de emisión del juicio solo hará uso de sus intuiciones morales; pero si es preciso hacer una justificación, entonces un proceso de razonamiento posterior vendrá a salvaguardar tanto a las intuiciones suscitadas como al juicio correspondiente —el razonamiento tiene, así, un carácter agregado—.

Haidt articula su modelo mediante la exposición de una serie de vínculos [*links*] que, según él, pueden surgir en la producción del juicio moral. En términos generales, los vínculos propuestos por el modelo son los siguientes: *i. El vínculo del juicio intuitivo*: en el que los juicios morales aparecen en la conciencia automáticamente y sin esfuerzo, como resultado de las intuiciones morales; *ii. El vínculo del razonamiento post hoc*: en donde el razonamiento moral sí resulta un proceso más bien esforzado, después de que el juicio moral se ha hecho y en el que la persona busca argumentos que apoyen sus juicios ya realizados; *iii. El vínculo de la persuasión razonada*: el razonamiento moral se

produce y se envía para exponer verbalmente a los demás la justificación de un juicio moral ya hecho; *iv. El vínculo de la persuasión social*: el cual, partiendo del hecho de que las personas están inmersas en un conjunto de normas, señala la influencia mutua y directa que se ejerce con el juicio moral entre amigos —aliados o conocidos—, aun cuando no se haya utilizado una persuasión razonada (cf. Haidt, 2001, 818-819).

Estos cuatro enlaces forman el núcleo del modelo intuicionista social; núcleo que si bien otorga al razonamiento moral un rol causal en el juicio moral, lo hace únicamente cuando el razonamiento se ejecuta a través de otras personas. Para Haidt, el papel de la razón es, entonces, secundario, puesto que las personas rara vez invalidan sus juicios intuitivos iniciales razonando en forma privada sobre sí mismas —rara vez, según Haidt, se utiliza el razonamiento para cuestionar las propias actitudes o creencias—. Sin embargo, aun cuando esto último no ocurra con frecuencia, es un hecho que ocurre, pues las personas están en la capacidad de participar en el razonamiento moral privado; muchas personas pueden señalar momentos de su vida en los que cambiaron de opinión sobre un problema moral con tan solo reflexionar sobre el asunto por sí mismos. Teniendo en cuenta este hecho, el modelo intuicionista social incluye dos maneras adicionales en las que el razonamiento privado puede formar los juicios morales.

Así, podemos hablar también de: *v. El vínculo del juicio razonado*: el cual señala que a veces las personas pueden razonar sus juicios por simple fuerza de la lógica, anulando así sus intuiciones iniciales (cf. Haidt, 2001, 819). En tales casos, el razonamiento es realmente causal y no puede decirse que sea “esclavo de las pasiones”. Sin embargo, Haidt encuentra que estos casos son raros y que ocurren, principalmente, en circunstancias en las cuales la intuición inicial es débil, mientras que la capacidad de razonamiento es alta. Además, él señala que en aquellos casos en los que el juicio

razonado entra en conflicto con un juicio intuitivo fuerte, las personas por lo general tienen una “actitud dualista”, puesto que el juicio razonado puede expresarse y, sin embargo, el juicio intuitivo continúa existiendo bajo la superficie.

Por último, Haidt complementa su modelo con un último vínculo: *vi. El vínculo de la reflexión privada*: de acuerdo con el cual, en el curso del pensamiento sobre una situación, una persona puede activar de manera espontánea una nueva intuición que contradice el juicio intuitivo inicial (cf. *ibíd.*). Con esto Haidt apunta al método más ampliamente discutido de desencadenar nuevas intuiciones, mediante la asunción de un rol, es decir, simplemente poniéndose uno mismo en los zapatos de otra persona. Asumir un rol nos puede permitir, instantáneamente, sentir dolor, simpatía u otras respuestas emocionales vicarias, puesto que si una persona llega a ver un problema o dilema desde más de una perspectiva, puede experimentar múltiples intuiciones en conflicto. Siendo así, el juicio final puede ser determinado de dos maneras, o bien siguiendo la intuición más fuerte, o bien permitiendo a la razón elegir entre las alternativas sobre la base de la aplicación consciente de una norma. Es pues esta la manera en que esta ruta se encamina a tener un diálogo interior con uno mismo, obviando la necesidad de un discurso acompañado.

Hasta aquí, podemos apreciar la forma en que el modelo intuicionista social de Haidt se apoya, principalmente, en los vínculos *i* a *iv* —puesto que son, en esencia, emocionales—, defendiéndolos como las maneras en que, normalmente, ocurren los juicios morales. Pero resulta claro que el modelo también admite que, en ciertos casos y por más extraños que resulten a la apreciación de Haidt<sup>1</sup>, estos juicios se pueden

---

<sup>1</sup> Para una exposición detallada de los problemas que Haidt encuentra en el modelo racionalista y que, según él, son solucionados por su modelo intuicionista —el problema del proceso dual, el problema del razonamiento motivado, el problema *post-hoc* y el problema de la acción—, véase Haidt, 2001, 819-825.

producir a partir del razonamiento, como se infiere de los vínculos *v* y *vi*. Sobre esto volveré al final de esta sección.

Pese a esto y para apoyar su modelo intuicionista social, Haidt ofrece un extenso conjunto de evidencia empírica, en el cual una de las más representativas es la indagación de las reacciones morales que tienen las personas ante determinadas situaciones. Así, por ejemplo, Haidt y sus colaboradores indagaron a un numeroso grupo de estudiantes acerca de su reacción ante un caso de incesto consensuado. El relato que se tomó como base fue el siguiente:

Julie y Mark son hermanos. Ellos viajan juntos a Francia, durante las vacaciones de verano de la universidad. Una noche, ellos se quedan solos en una cabaña cerca a la playa. Allí, ellos deciden que sería interesante y divertido si intentaran hacer el amor; por lo menos, sería una nueva experiencia para cada uno de ellos. Julie ya estaba tomando pastillas anticonceptivas, pero Mark decide utilizar también un preservativo, sólo para estar seguro. Ambos disfrutaban haciendo el amor, pero deciden no volverlo a hacer. Ellos mantienen esa noche como un secreto especial, que les hace sentirse mucho más unidos. ¿Qué piensa usted de esto? ¿Estuvo bien que ellos hicieran el amor? (Haidt, 2001, 814).

La mayoría de las personas que consideraron la situación la valoraron como moralmente incorrecta, pero tuvieron grandes dificultades para explicar por qué. Cada vez que surgía un argumento para demostrar que los hermanos habían hecho algo inmoral, los investigadores daban razones de por qué el argumento no era satisfactorio (cf. Haidt, Bjorklund, & Murphy, 2000). Por ejemplo, muchas personas se preocupaban por la concepción de un niño deformado, pero ante este temor los investigadores recordaban que se habían utilizado anticonceptivos. Algunos estaban preocupados por los efectos sobre la comunidad, pero esto no se aplicaba porque Julie y Mark no se lo habían contado a nadie. Otras personas señalaban que los jóvenes podrían haber quedado traumatizados, pero los investigadores recordaban que ellos realmente habían disfrutado de la experiencia y que esta fortaleció su relación. Incluso, otras personas trajeron a colación que el incesto es condenado en la Biblia, pero no

podieron recordar en qué parte. De esta manera, a cada argumento que señalaba al acto como inmoral se respondió con un contraargumento. Sin embargo, pese a que las personas indagadas reconocieron que los contraargumentos eran contundentes, solo el 17% cambió su juicio moral inicial; los demás, apelaron a exclamaciones emocionales, tales como ¡El incesto es desagradable!, ¡El incesto no solo es malo: es salvaje! En consecuencia, aunque no hubo razones como apoyo —en tanto que muchos se remitían a decir: “Yo no sé... Yo no puedo explicar esto... Yo sólo sé que es malo” (Haidt, 2001, 814)—, la valoración moral, fundamentada en emociones negativas, se mantuvo en un rechazo rotundo. La conclusión que extrae Haidt de este fenómeno, a la cual denomina “moral sin habla” [*moral dumbfounding*], es que el razonamiento normalmente no desempeña ningún papel en la producción del juicio moral.

Este tipo de evidencia apunta a demostrar no solo que las emociones pueden ejercer una fuerte influencia en los juicios morales, sino también que ellas son tanto necesarias como suficientes para este tipo de juicios. Sin embargo, pienso que aquí cabría preguntarse si realmente el ejemplo en cuestión constituye un caso de juicio moral. Si bien es cierto que se invita a las personas a valorar la situación desde una perspectiva moral: por una parte, no resulta del todo claro que haya aquí un agente moral y mucho menos una persona afectada que den lugar a un juicio moral; por otra parte, a partir de las respuestas de los estudiantes indagados, bien podríamos sugerir que estas respuestas constituyen una valoración de ‘gusto’, mas no una valoración moral —el encuestado puede interpretar la pregunta como si esta dijera: ¿le gustaría a usted hacer lo mismo que hicieron Julie y Mark?—. A partir de este tipo de evidencia, inferir la necesidad y suficiencia de las emociones en los juicios morales no es, entonces, del todo concluyente.

Con todo, Haidt considera que su conclusión es muy fuerte y que se apoya, además, en la investigación sobre los efectos de la inducción de la emoción. Así, por ejemplo, Haidt y Wheatley (Wheatley y Haidt, 2005) hicieron después un estudio en el cual hipnotizaron personas para que sintieran una punzada de disgusto al escuchar las palabras, emocionalmente neutrales, “tomar” o “a menudo”. Luego se les pidió que valoraran moralmente al protagonista de varios relatos que contenían una de estas dos palabras, pero unos de los cuales eran moralmente problemáticos y otros sin relevancia moral alguna. A la mitad de los participantes se les dio versiones de los relatos que incluían la palabra inducida hipnóticamente, mientras que la otra mitad recibió versiones casi idénticas, solo que sin la palabra en cuestión. Este es uno de los relatos moralmente problemáticos:

El congresista Arnold Paxton con frecuencia da discursos condenando la corrupción y abogando por la reforma de las finanzas de campaña. Pero él sólo está tratando de encubrir el hecho de que él mismo [**toma** sobornos de / es **a menudo** sobornado por] las tabacaleras, y otros intereses especiales, a fin de promover su legislación (Wheatley y Haidt, 2005, 781).

Y este es uno de los relatos moralmente neutrales:

Dan es un representante en el consejo estudiantil de su colegio. Este semestre él está a cargo de la programación de debates sobre cuestiones académicas. Él [intenta **tomar** / **a menudo** recoge] temas que atraen tanto a profesores como estudiantes con el fin de estimular el debate (Wheatley y Haidt, 2005, 782).

La indagación arrojó como resultado que los participantes juzgaran al protagonista de las situaciones moralmente problemáticas como obrando de la peor manera posible, pero la presencia de la palabra inducida hipnóticamente hizo que, también, se condenara moralmente a los protagonistas de las situaciones no problemáticas. De aquí que el personaje Dan, haya sido juzgado como moralmente reprochable por los participantes, aunque no podían decir por qué; ellos solo se remitían a expresar: “Aquí hay algo sospechoso”. Es pues en este sentido que, dejando a



la razón sin habla, Haidt apoya su tesis de que el razonamiento normalmente no desempeña ningún papel en la producción del juicio moral; tesis que, sin embargo y como lo señalé anteriormente, podemos develar como no del todo concluyente.

Pero el hecho de que Haidt acepte que esto, “normalmente”, sucede así, no es un aspecto nada despreciable en el debate en cuestión. Esto resulta importante porque el emocionalismo de Haidt no es, entonces, excluyente con la postura racionalista, sino que tan solo le da una mayor preeminencia a las emociones como fuentes principales de los juicios morales. Dos observaciones considero que es pertinente hacer entonces a este respecto, las cuales se conectan con la conclusión que obtuvimos de los vínculos expuestos por Haidt. Por una parte, pienso que aceptar que el razonamiento pueda producir juicios morales, sin la ayuda de las intuiciones, hace que de la postura de Haidt no se pueda inferir un emocionalismo fuerte. Por otra parte, debemos tener en cuenta que el autor también acepta que, pese a la evidencia que apoya la existencia de los vínculos *i* a *iv* en algunos ámbitos del juicio, los vínculos del modelo no necesariamente se deben dar en el ámbito del juicio moral; de aquí que, según él, su modelo se presente más bien como “[...] una propuesta para estimular el pensamiento y las nuevas investigaciones sobre los juicios morales” (Haidt, 2001, 818). Pienso que estos dos factores, la aceptación de la existencia de los vínculos racionales y la aceptación de la no necesidad de los vínculos emocionales, hacen entonces que la postura de Haidt ya no sea del todo emocionalista. De aquí que, como lo sugerí anteriormente, esta postura puede ser caracterizada como un emocionalismo más bien débil, esto es, como una postura que le da una mayor preeminencia a las emociones como fuentes principales de los juicios morales sin negar que el razonamiento también pueda ser una de estas fuentes. En este sentido, es preciso observar que en la postura de Haidt el debate entre emocionalistas y racionalistas aún no incurre en un radicalismo unilateral.

### 2.1.2 El emocionalismo de Jesse J. Prinz

Como lo sugerí anteriormente, la postura emocionalista cobra toda su fuerza en la filosofía de Jesse J. Prinz, puesto que y como veremos a continuación, es él quien expone este enfoque de la manera más sistemática, completa, coherente y fundamentada (desarrollando *in extenso* las tesis de Hume con respecto a la construcción emocional de la moralidad). Pienso que, en esencia, Prinz está completamente de acuerdo con Haidt, pues él encuentra que el trabajo de este último, junto con otros trabajos recientes en ciencia cognitiva, proveen una evidencia arrolladora para establecer una relación entre la emoción y el juicio moral. El trabajo de Prinz parte, en buena medida, del trabajo de Haidt y de aquí la pertinencia de haber expuesto primero la postura de este último.

Es así como revisando hallazgos de psicología, neurociencia cognitiva, investigación sobre psicopatología y algunas observaciones antropológicas, en un primer momento Prinz defendió que “[...] las emociones no están simplemente relacionadas con los juicios morales sino que ellas también son, en algún sentido, tanto necesarias como suficientes para dichos juicios” (Prinz, 2006, 29). He dicho “en un primer momento”, porque Prinz aclara la expresión ‘en algún sentido’ en una corrección posterior a su postura, tal y como podemos ver en los dos siguientes pasajes:

La división entre los teóricos que piensan que los sentimientos son esenciales a la moral y aquellos que piensan que las emociones son incidentales es quizás la ruptura más fundamental en la filosofía moral. Yo estoy del lado de los miembros del primer grupo [...] Yo usaré el término “emocionalismo” [*emotionism*] como una etiqueta para cualquier teoría que diga que las emociones son de algún modo esenciales (Prinz, 2007a, 13).

Yo dije que los emocionalistas postulaban una “relación esencial” entre las emociones y las cosas en el dominio moral. ¿Qué es lo que está esencialmente relacionado? Elijo esta frase, en lugar de “necesariamente relacionado” porque puede haber algo de libertad entre la necesidad y la esencia. Algo *A* pertenece a la esencia de otra cosa *B* si uno no puede especificar qué es *B* sin mencionar a *A*. Esta formulación no invoca la necesidad en un sentido modal fuerte. No

dice que todos los *B* son necesariamente *A*. Uno *puede* construir relaciones esenciales en este sentido fuerte (Prinz, 2007a, 19).

A esta relación esencial, que Prinz encuentra finalmente entre las emociones y los juicios morales, él la denomina ‘sentimentalismo constructivo’ o ‘emocionalismo’. Se trata de una forma de teoría de la sensibilidad, de acuerdo con la cual los hechos morales son productos de nuestras reglas morales; reglas que, a su vez, están constituidas por sentimientos, establecidos mediante interacciones bioculturales. De aquí que Prinz vaya mucho más allá que Haidt, al defender que no podemos hablar de moralidad sin hacer referencia a las emociones.

En términos generales, es de anotar la concepción que de las emociones se encuentra a la base de esta postura. Prinz desarrolla una teoría de las emociones, que se construye a partir de la concepción somática de James-Lange. Él complementa esta concepción con la teoría de la apreciación encarnada [*embodied appraisal theory*] para hacer frente a algunas objeciones y concluir, así, que si bien las emociones son señales somáticas o cambios corporales, también son apreciaciones que representan preocupaciones [*concerns*]. Esta representación, sin embargo, es no-conceptual y, en consecuencia, las emociones no son estados cognitivos. Y esto implica, de acuerdo con Prinz, que:

[...] una emoción puede estar garantizada o no garantizada [*unwarranted or warranted*], y esto también permite que dos emociones puedan tener diferente contenido a pesar de ser somáticamente indistinguibles. [...] yo he argüido que las teorías cognitivas están en lo cierto acerca del contenido de las emociones, pero las no-cognitivas están en lo cierto acerca de la forma (ibíd. 68).

Muchas objeciones se podrían plantear a esta concepción ‘híbrida’ de las emociones —en particular la que apunta a cuestionar el contenido representacional que de la emoción defiende Prinz—, pero un examen detallado de estas objeciones desbordaría los límites y propósitos de este trabajo. Solo cabría agregar la manera en

que, según Prinz, las emociones pueden dar lugar a nuevas emociones y, a su vez, pueden llegar a convertirse en emociones morales. Para el autor, las emociones están asociadas con archivos mentales que las calibran a preocupaciones; en esta recalibración se establece un nuevo archivo mental, que relaciona la emoción existente con un nuevo conjunto de condiciones suscitadoras [*elicitors*] (cf. *ibíd.*, 67). Esto explica cómo puede surgir una emoción frente a nuevas situaciones y, así mismo, cómo puede dar lugar a nuevas emociones. Además, siendo que estos archivos mentales pueden constituir reglas morales, Prinz complementa este proceso señalando que una emoción moral es, entonces, “[...] una emoción que ocurre en respuesta a algo que es acorde o falla en estar acorde con una regla moral” (*ibíd.*, 118).

Así pues, Prinz usa la evidencia empírica como soporte para la defensa de varias tesis filosóficas; tesis que, en su conjunto, pueden verse como una afirmación de la tesis emocionalista en su vertiente moral. En primer lugar, el autor defiende que el sentimentalismo es verdadero y que, en consecuencia, juzgar que algo es erróneo es tener un sentimiento de desaprobación al respecto; en segundo lugar, afirma que los hechos morales son respuestas-dependientes: lo malo es justamente lo que causa desaprobación en una comunidad moral. Y, en tercer lugar, Prinz se compromete con la defensa de que una forma de internalismo motivacional es verdadero, esto es, que los juicios morales ordinarios están intrínsecamente motivados y que todos los juicios morales no-motivados son dependientes de los primeros (cf. *ibíd.*, 19).

En lo que sigue, para la exposición de la primera tesis me apoyaré en un artículo de Prinz, titulado “The Emotional Basis of Moral Judgments” (cf. Prinz, 2006), el cual contiene, en esencia, los planteamientos básicos del emocionalismo moral. Para la exposición de las dos tesis restantes, sin embargo, considero que es mejor apelar a la manera en que estas se presentan en el último libro del autor, titulado *The Emotional*

*Construction of Morals* (cf. Prinz, 2007a), debido a una mayor consistencia con la que allí son presentadas. Dicho esto, veamos entonces cómo articula Prinz la defensa de su tesis emocionalista con la cual pretende llevar a cabo una construcción emocional de la moral.

### 2.1.2.1 *Las bases emocionales del juicio moral*

#### 2.1.2.1.1 *Co-existencia de la emoción en el juicio moral*

Prinz parte de la siguiente pregunta: “¿tienen nuestros conceptos morales *comunes* un componente emocional (aquellos que, con mayor frecuencia, desplegamos como guía del pensamiento)?” (Prinz, 2006, 30). Para él, esta es, en esencia, una cuestión empírica, que indaga por lo que ocurre en nosotros cuando usamos términos morales tales como “bueno” y “malo” o “correcto” e “incorrecto”. De aquí que si la evidencia empírica responde de manera afirmativa a la pregunta formulada, esto será una fuerte prueba de que las emociones co-existen con los juicios morales.

El primer argumento (cf. Prinz, 2006, 30-31) señala la manera en que —a partir de diferentes tipos de investigaciones— se ha encontrado evidencia a favor de la activación de las áreas emocionales del cerebro cuando las personas hacen juicios morales. Así, por ejemplo, Moll, de Oliveira-Souza, y Eslinger (2003) midieron la actividad cerebral en un grupo de personas mediante la evaluación de proposiciones morales tales como “Usted debe violar la ley cuando sea necesario”, en contraste con proposiciones fácticas tales como “Las piedras están hechas de agua”. En ambos casos, las personas simplemente tenían que responder “correcto” o “incorrecto”. Ellos descubrieron que cuando las personas hicieron juicios morales, en contraste a cuando hicieron juicios fácticos, las áreas del cerebro que están relacionadas con la reacción emocional estuvieron activas. En un estudio diferente, (cf. Sanfey et al., 2003) se midió

la actividad cerebral en personas mediante un ‘*ultimatum game*’<sup>2</sup>. En cada caso, se le pidió a un jugador que dividiera una cantidad de dinero con otro jugador. Cuando la división fue demasiado injusta, el segundo jugador tuvo actividad cerebral en áreas relacionadas con la emoción. Por otro lado, Berthoz et al. (2002) encontraron enlaces similares con áreas cerebrales de la emoción cuando las personas consideraron violaciones de reglas sociales. En este caso, a las personas se les contó una historia de un invitado a cenar, quien, después de probar la comida, la espetó groseramente en una servilleta, sin disculparse. De igual manera, Prinz trae a colación los hallazgos de Greene et al. (2001) —que veremos más adelante con mayor detalle—, quienes encontraron activación emocional en personas que consideraron dilemas morales. Estos trabajos son comparables también a los de Kaplan, Freedman, y Lacoboni (de próxima aparición), quienes encontraron activación emocional cuando las personas miraron fotografías de políticos a quienes ellos se oponían.

#### 2.1.2.1.2 Suficiencia de la emoción para el juicio moral

La evidencia es muy sugestiva, pero el mismo autor es consciente de que no es del todo concluyente, pues si bien es cierto que este tipo de evidencia señala que las emociones co-ocurren con los juicios morales, no se sigue de aquí que las emociones estén involucradas con estos juicios de una manera directa. De aquí que una nueva pregunta sea, entonces: ¿Son las emociones efectos simples de los juicios morales o están involucradas de manera más primordial? Para responder a esta pregunta, Prinz apela a otro tipo de evidencia, con la cual apuntará a defender, a su vez, una segunda tesis: “[...] las emociones influyen en los juicios morales” (Prinz, 2006, 31).

---

<sup>2</sup> En este juego participan dos sujetos, a los que se ofrece una buena cantidad de dinero que podrán repartirse entre ellos si se ponen de acuerdo en la repartición. A uno se le dará la facultad de proponer el trato y el otro sólo tendrá dos opciones, aceptar o rechazar la oferta. Si la acepta, cada uno se llevará la parte acordada. Si la rechaza, ninguno recibirá nada.

La evidencia es suministrada ahora por Schnall, Haidt, y Clore (de próxima aparición), quienes dieron a un grupo de personas un cuestionario con una serie de viñetas y les pidieron que evaluaran lo malo o incorrecto de las acciones descritas. En estas viñetas, las personas leyeron, por ejemplo, lo siguiente:

Al perro de Frank lo mató un automóvil en frente de su casa, así que él cortó el cuerpo, lo cocinó y se lo comió en la cena. ¿Qué tan malo es esto?

Es de anotar que la mitad de las personas que leyeron estas viñetas estaban sentadas en un escritorio limpio y bonito, mientras que la otra mitad estaba sentada en un escritorio sucio, con una taza de bebida crujiente, un lápiz mascado, un pañuelo de papel usado y una caja de pizza grasienta. La observación obedece a que, de acuerdo con los resultados, las personas en el escritorio repugnante evaluaron las viñetas como más malas que las personas en el escritorio limpio. Así, la segunda tesis, esto es, que las emociones influyen en los juicios morales, se ve apoyada por este ejemplo, en tanto que una emoción negativa —a raíz de percepciones desagradables— puede llevarnos a hacer valoraciones morales más negativas que las que podríamos hacer en su ausencia.

Otra evidencia, que Prinz trae a favor de esta tesis, son los dos experimentos realizados por Haidt y su equipo de trabajo, con respecto a las personas hipnotizadas para que sintieran una punzada de disgusto al escuchar determinadas palabras y las respuestas ante el caso del incesto, tal y como los vimos anteriormente (cf. *infra*, secc. 2.1.1). De manera similar a como lo hizo Haidt, Prinz concluye que este ejemplo es una evidencia fuerte a favor de que un sentimiento negativo puede fomentar una valoración moral negativa; esto sin que medie ningún tipo de creencia específica sobre alguna propiedad en virtud de la cual algo sea malo. En consecuencia, tales conclusiones sugieren que podemos construir la creencia de que algo es moralmente malo con tan solo tener una emoción negativa dirigida hacia eso.

### 2.1.2.1.3 Necesidad de la emoción para el juicio moral

Hasta aquí, sin embargo, tenemos que las emociones parecen ser suficientes para una valoración moral, pero ¿serán acaso también necesarias? A este respecto, Prinz apela al papel fundamental que juegan las emociones en el desarrollo moral, exponiendo la manera en que un niño aprende la moralidad mediante la educación emocional (cf. Prinz, 2006, 32). Sin embargo, Prinz cree que el argumento fundamental para apoyar el carácter necesario de las emociones en el desarrollo moral deriva de la investigación acerca de los psicópatas. En su opinión:

Los psicópatas son la prueba fehaciente de la tesis de la necesidad, porque son extremadamente deficientes para tener emociones negativas, como el miedo y la tristeza, en particular. Rara vez, ellos experimentan estas emociones y tienen una gran dificultad para reconocerlas, siquiera en las expresiones faciales y los sonidos del habla (Blair et al., 2001, 2002). Los psicópatas no son sensibles a las condiciones de miedo, experimentan el dolor con una intensidad menor que los sujetos normales y no se perturban con fotografías que nos causan sufrimiento (Blair et al., 1997). Esto indica que la psicopatía resulta de un déficit de baja intensidad en las emociones negativas. Sin un núcleo de emociones negativas, no pueden adquirir el sufrimiento empático, el remordimiento o la culpa. Estos déficits emocionales parecen ser la causa original de sus patrones de comportamiento antisocial. Pienso que los psicópatas actúan de mala manera porque no pueden hacer juicios morales genuinos. Dan la aprobación a la moral solo de dientes para afuera, pero hay una buena razón para pensar que no tienen conceptos morales —o por lo menos no tienen los conceptos morales como las personas normales—. Los psicópatas reconocen que sus actos delictivos son “malos” pero no comprenden la importancia de esta palabra (Prinz, 2006, 32).

De esta forma y apoyándose ahora en Cleckley (1941), Prinz está de acuerdo en que un psicópata es como un daltónico, puesto que el psicópata puede decir que comprende lo bueno y lo malo, pero no hay ninguna manera para él de darse cuenta de que no comprende. En otra investigación que Prinz trae a colación, Blair (1995), se investigó el rol de los conceptos morales en los psicópatas y se descubrió que ellos trataban las injusticias morales como si, simplemente, fueran convencionales. En este sentido, los psicópatas entienden el término “malo” o “incorrecto” como si tan solo significara “prohibido por las autoridades locales” (cf. *ibíd.*).



En suma, Prinz defiende que los psicópatas fallan en la comprensión de los conceptos morales, sin importar su alto nivel de inteligencia, pues ellos pueden aprender todo lo que hay por aprender sobre las consecuencias de una acción sin entender que esta es inmoral. De aquí que la investigación sobre psicopatía sugiera que las emociones son, en cuanto al desarrollo, necesarias para adquirir la capacidad de elaborar juicios morales.

En mi opinión, este argumento no es del todo concluyente, pues tal y como Prinz lo presenta parece incurrir en una petición de principio y, además, en una generalización apresurada. Por una parte, cuando Prinz afirma que “los psicópatas fallan en la comprensión de los conceptos morales [...]”, pues ellos pueden aprender todo lo que hay por aprender sobre las consecuencias de una acción sin entender que esta es inmoral”, presupone, por supuesto, lo que se esperaría que pruebe, esto es, que entender que una acción es inmoral implica tener sentimientos negativos hacia esa acción. Por otra parte, es preciso tener en cuenta que una cosa es que los psicópatas carezcan de algunas emociones negativas y otra, bien diferente, que carezcan ‘del todo’ de emociones negativas. Así, para utilizar un ejemplo del mismo autor, considérese el caso de los Guhuku-Gama de Nueva Guinea y otros caza-cabezas, a quienes no consideramos como psicópatas. Ellos creen que es moralmente incorrecto matar a un miembro de su grupo familiar, pero que está perfectamente bien matar a otros; también piensan que es moralmente correcto que otros maten a sus familiares y que está bien matar a personas inocentes. Los Guhuku-Gama no tienen, por tanto, sentimientos negativos hacia la acción de matar a personas inocentes, pero en tanto que tienen sentimientos negativos hacia el hecho de matar a un miembro de su grupo familiar sí son individuos morales. Un caso similar es el que representan los psicópatas, pues su incapacidad para sentir empatía o culpa está dirigida hacia un determinado

grupo de personas o conjunto de acciones, pues bien podrían asesinar personas adultas pero considerar inadmisible asesinar a niños, por ejemplo.

Por otro lado, un segundo argumento que Prinz trae como apoyo de la necesidad de las emociones en el juicio moral es la que él denomina una tesis disposicional (cf. Prinz, 2006, 32). De acuerdo con esta tesis, podemos decir cosas tales como “El homicidio es malo”, sin sentir ninguna emoción, de manera similar a como decimos que los plátanos son amarillos, sin formarnos una imagen mental del amarillo. Pero, pregunta Prinz, ¿siendo sinceros, podemos atestiguar que asesinar es moralmente malo sin estar dispuestos a tener emociones negativas hacia el homicidio? En su opinión, una persona que responda afirmativamente a esta pregunta podría estar confundida o, realmente, no estaría siendo sincera. Para respaldar esta intuición, Prinz propone un experimento mental. Él nos invita a imaginar a una persona que sabe todo lo no-emocional relacionado con el homicidio. Esta persona sabe que el homicidio disminuye la utilidad y que sería casi irracional universalizar la máxima “Usted debe matar”. No podríamos decir que esta persona cree que el homicidio es malo, porque bien podría ser que ella creyera en todo lo que sabe acerca del homicidio sin tener ninguna postura moral sobre el mismo o, incluso, ninguna comprensión de lo que representaría decir que el homicidio es malo. Y, en cambio, señala Prinz, si una persona albergara un fuerte sentimiento negativo hacia el homicidio, diríamos que cree que el homicidio es moralmente malo, aun cuando no tuviera ninguna creencia explícita sobre si el homicidio disminuye la utilidad o si resulta contradictorio con la voluntad general.

El tercer y último argumento de Prinz, a favor de la necesidad de las emociones en el juicio moral, es uno de tipo antropológico. Este argumento, sencillamente, apunta a traer a colación numerosos ejemplos de la alta diversidad moral existente en el

mundo, tanto desde un punto de vista espacial como a través del tiempo. Así, si los juicios morales estuvieran basados en algo fuera de las emociones —en algo así como la razón o la reflexión— esperaríamos más convergencia moral transculturalmente, puesto que razón y reflexión convergen por encima del tiempo. Sin embargo, señala Prinz, la evidencia —la manera en que es el mundo— señala todo lo contrario:

El Guhuku-Gama de Nueva Guinea y otros caza-cabezas piensan que está bien matar a personas inocentes; los ciudadanos griegos de Tolomeo, en Egipto, se casan con sus hermanos en una tasa de hasta el 30%; los aztecas de México e incontables sociedades de pequeña escala se permitieron el canibalismo; los romanos llenaron ruedos para observar a gladiadores masacrarse; los hombres de Thonga tienen relaciones sexuales con sus hijas antes de cazar; las mujeres de China soportaron dolor terrible amarrando sus pies; la inequidad de género y la esclavitud han sido ampliamente aceptadas y ampliamente condenadas. Más cerca de casa, encontramos los debates interminables entre liberales y conservadores. También encontramos las diferencias regionales: los hombres blancos del sur son mucho más propensos, que sus homólogos del norte, a aprobar moralmente las represalias violentas para los insultos públicos, y otras ofensas no violentas. Estos ejemplos no son exóticos. Cualquier par de culturas elegidas al azar tendrán diferencias dramáticas en valores morales, y muchas de estas diferencias (como la poligamia vs. la monogamia o la belicosidad del sur vs. la diplomacia del norte) no tienen ninguna base en las diferentes creencias factuales. (Prinz, 2006, 33).

Esto sugiere, entonces, que los valores morales básicos no tienen un origen simplemente cognitivo. De acuerdo con Prinz, pese a que esta divergencia moral no demuestra directamente que las emociones son un componente necesario de la moral, sí provee una evidencia indirecta para apoyar esta conclusión. Así, si los valores morales no son conducidos por la razón o la reflexión, entonces es posible pensar que dependen de las pasiones culturalmente inculcadas.

### 2.1.2.2 *Una teoría sentimental del juicio moral*

Ahora bien, pese a toda la argumentación que Prinz ha defendido hasta aquí, él mismo señala que ninguna de estas pruebas empíricas mencionadas suministra un argumento demostrativo para una teoría del juicio moral sustentada en las emociones. El autor es consciente de que bien podría ser el caso que los juicios morales estuvieran

correlacionados y causalmente relacionados con respuestas emocionales sin involucrar estas repuestas de manera esencial. Es por ello que él considera necesario exponer una teoría del juicio moral que sistematice los datos de los que ha estado hablando, una teoría que, a su parecer, “[...] ofrece una explicación y quizás una mejor explicación que la de muchas otras versiones” (ibíd.). Dicha teoría es una variante de la postura de Hume, según la cual: “Creer que algo es moralmente incorrecto (correcto) es tener un sentimiento de desaprobación (aprobación) hacia eso” (ibíd.).

Prinz utilizará aquí el término ‘sentimiento’ como una disposición a tener emociones. Su idea es defender una tesis sentimentalista, según la cual cuando determinamos que algo es incorrecto surge un sentimiento de desaprobación que, a su vez, puede estar constituido por diferentes emociones. De esta manera, una u otra emoción se despertará frente a la acción considerada como incorrecta y el juicio resultante será, entonces, una expresión de la disposición emocional o sentimiento subyacente. En este sentido, un juicio de que algo es incorrecto consistirá en la disposición respectiva y este juicio contendrá normalmente una emoción específica que manifiesta esta disposición (cf. Prinz, 2006, 34). Así, por ejemplo, una desaprobación puede despertar un sentimiento de culpa, pero hay varias emociones diferentes en esta categoría. Cuál de estas emociones experimentemos depende de quién esté siendo culpado y por qué, ya que si hacemos algo malo, podemos experimentar vergüenza o culpa, mientras que si otra persona hace algo malo, podemos experimentar ira, desprecio o aversión. El punto principal de Prinz aquí es que, en cualquier caso, el juicio resultante es la expresión del sentimiento de culpa, matizado o graduado de acuerdo con la emoción que se haya despertado; emoción que, entonces, puede verse como una categoría que se extiende a lo largo de todo un rango de subtipos, los cuales miden el grado de desaprobación.

Es en este sentido en que Prinz afirma que: “[...] la emoción sirve como un vehículo del concepto “incorrecto” en la misma forma en que una idea de algún color específico podría servir como un vehículo para el pensamiento de que las cerezas son rojas. Las instancias [*tokens*] del concepto “incorrecto” pueden ser, así, idénticas a las emociones” (ibíd.). Es así como Prinz concluye la defensa de la primera parte de su tesis emocionalista moral, esto es, que el sentimentalismo es verdadero y que, en consecuencia, juzgar que algo es erróneo es tener un sentimiento de desaprobación al respecto.

### 2.1.2.3 *Emocionalismo fuerte: emocionalismo metafísico y epistemológico*

Las segunda y tercera partes de la tesis emocionalista moral dicen que, por una parte, los hechos morales son respuestas-dependientes: lo malo es justamente lo que causa desaprobación en una comunidad moral; y, por otra parte, que la tesis emocionalista moral se compromete con la defensa de que una forma de internalismo motivacional es verdadera, esto es, que los juicios morales ordinarios están intrínsecamente motivados y que todos los juicios morales no-motivados son dependientes de los primeros.

Comencemos con la segunda parte. Para Prinz, los juicios morales expresan sentimientos y los sentimientos se refieren a la propiedad o disposición de causar ciertas reacciones en nosotros; estas reacciones, a su vez, él las concibe como “[...] sentimientos de cambios corporalmente moldeados [*patterned bodily changes*]” (ibíd.). A partir de aquí, la idea de Prinz es defender entonces una teoría de la sensibilidad, de acuerdo con la cual los conceptos morales se refieren a propiedades de respuesta-dependiente (cf. Dreier 1990; Johnston 1990; McDowell 1985; McNaughton 1988; Wiggins 1991; y Wright 1992). Estas propiedades de respuesta-dependiente a las que se refieren los sentimientos —y, por extensión, los conceptos morales— podrían

ejemplificarse cuando decimos que el chocolate es rico, puesto que aquí atribuimos al chocolate la propiedad de causar, por ejemplo, placer en nosotros. Los juicios morales son, así, aptos-a-la-verdad<sup>3</sup> [*truth-apt*], en tanto que ellos hacen referencia a las propiedades de respuesta-dependiente, como acabamos de indicar.

Como complemento de esta idea —y aquí viene la tercera parte de la tesis—, de acuerdo con Prinz, no importa si con dichas propiedades estamos haciendo referencia a algo intrínseco o relacional a los hechos morales, sino que de cualquier manera suponemos que ellas dependen de las reacciones que estos hechos causan en nosotros. Es este el sentido en el que Prinz defiende que una forma de internalismo motivacional es verdadera, pues los juicios morales deben estar intrínsecamente motivados. Pero si esto es así, si lo ‘incorrecto’ se refiere a una propiedad de respuesta-dependiente, la pregunta que surge es: ¿cuáles respuestas importan? ¿Bajo qué condiciones? Previendo esta cuestión, Prinz acoge el relativismo-del-hablante<sup>4</sup> [*speaker-relativism*] de Dreier. Para Prinz, cuando decimos que algo es incorrecto, hacemos referencia —quizás inconscientemente— a la propiedad de causar emociones de culpa en nosotros —o, quizás, decir que algo es incorrecto quiere decir que esto causa emociones de culpa en nosotros, donde el “nosotros” se refiere a un grupo de personas con quien podríamos disentir—.

Desde esta perspectiva, los conceptos morales comunes no se refieren a las propiedades que pueden ser caracterizadas coherentemente sin hacer referencia a

---

<sup>3</sup> Una oración es apta-a-la-verdad [*truth-apt*] si hay algún contexto en el cual puede ser pronunciada [con su significado actual] y exprese una proposición verdadera o falsa. Las oraciones que no son aptas a la verdad incluyen preguntas y órdenes y, más controversialmente, oraciones paradójicas de la forma de la paradoja del mentiroso (‘esta oración es falsa’); u oraciones (‘usted no debería fumar’) cuya función aparente es hacer una afirmación, pero que en cambio pueden ser consideradas como la expresión de prescripciones o actitudes, en lugar de pretender aspirar a la verdad o a la falsedad.

<sup>4</sup> El relativismo del hablante [*speaker-relativism*] es la teoría según la cual el contenido de (o lo que es expresado por) una oración que contiene un término moral varía de acuerdo con (en función de) el contexto en el cual es usada. Cf. Dreier, James. “Internalism and Speaker Relativism”. En: *Ethics*, 101, Universidad de Chicago: Octubre, 1990, 6-26, pág. 6.

nuestros sentimientos. Para Prinz, no existe ningún concepto que se co-refiera a nuestros sentimientos de tal manera que o bien no contenga o bien no anuncie nuestros sentimientos —cualquier concepto que es co-referencial a nuestros sentimientos morales es, en este sentido, parásito de estos sentimientos—; de aquí que, en su opinión:

Los éticos normativos introducen conceptos que se refieren a algo aparte de aquello a lo que nuestros conceptos morales corrientes se refieren. ¿Existe alguna razón para llamar a los conceptos que ellos introducen “morales”? yo pienso que esta es una pregunta sin una respuesta definida. Los éticos normativos introducen conceptos que, de la misma manera que los conceptos morales comunes, son diseñados para regular el comportamiento, pero estos conceptos son diferentes de los conceptos que normalmente expresamos cuando usamos palabras tales como “correcto” o “incorrecto”. Si nosotros llamamos o no a tales conceptos “morales” es una cuestión de elección. El punto sobre el que quiero enfatizar es que los conceptos que normalmente expresamos usando un vocabulario moral están relacionados esencialmente con nuestros sentimientos y, en este sentido, el internalismo motivacional es verdadero (Prinz, 2006, 41).

Es de esta forma como Prinz llega a la conclusión de que albergar una creencia moral es tener un sentimiento de aprobación o de desaprobación, conclusión que, en su opinión, hace la mayor justicia a la evidencia aportada aquí. Así pues, su idea es que las emociones co-ocurren con los juicios morales, influyen en los juicios morales, son suficientes para los juicios morales y son necesarios para los juicios morales, porque los juicios morales están constituidos por disposiciones emocionales. Por lo menos, de acuerdo con esta postura, nuestros conceptos morales corrientes parecen tener esta característica.

En suma, Prinz presenta una revisión de las diferentes formas en las cuales las emociones pueden estar implicadas en la moralidad. Esta revisión lo lleva a introducir el término ‘emocionalismo’ [*Emotionism*] para etiquetar cualquier punto de vista que considere dicha implicación como algo esencial; aunque, valga decirlo, el mismo autor confiesa que los argumentos que apoyan esta conclusión no son del todo concluyentes.

Es por esto que Prinz termina defendiendo su postura como un argumento a la mejor explicación, la cual se sistematiza en la que —como vimos— él denomina una teoría sentimental del juicio moral. El emocionalismo es pues la tesis que, como argumento a la mejor explicación, afirma que la moral está relacionada con las emociones de manera esencial (Prinz, 2007a, 13). Prinz distingue entre dos clases de emocionalismo: metafísico y epistemológico. El ‘emocionalismo metafísico’ es la tesis de acuerdo con la cual las propiedades morales están esencialmente relacionadas con las emociones (Prinz, 2007a, 14), y el ‘emocionalismo epistemológico’ es la tesis que afirma que los conceptos morales están esencialmente relacionados con las emociones (Prinz, 2007a, 16). Prinz apoya, entonces, ambas clases de emocionalismo y a esta conjunción la denomina “Emocionalismo fuerte”.

## 2.2 EL INNATISMO MORAL

A pesar de los hallazgos de Haidt y de la muy coherente y fundamentada sistematización del emocionalismo que hace Prinz, algunos rechazan la idea de que una estructura emocional ocupe el primer lugar en la producción del juicio moral. Entre ellos, se encuentran los defensores del innatismo moral (Dwyer, 1999 y Hauser, 2006, 2008, principalmente), postura que afirma que venimos equipados con una especie de facultad moral que opera sobre las propiedades causales-intencionales de eventos y acciones moralmente relevantes. A su vez, esta postura se apoya en la lingüística de Noam Chomsky, que defiende la existencia de un sistema de principios universales en nuestra mente (Gramática Universal) —que es necesario para el desarrollo del lenguaje—. Además, el innatismo moral también hace eco de la analogía que John Rawls, en un pasaje de la *Teoría de la Justicia* y apoyándose en la lingüística de Noam Chomsky, estableció entre el lenguaje y la moral. Allí el autor escribe lo siguiente:



Una comparación muy útil aquí se refiere al problema de describir el sentido de la gramática del cual nos servimos para construir las oraciones de nuestra lengua materna\*. En este caso, el objetivo es caracterizar la capacidad de reconocer oraciones bien formadas mediante la formulación de principios claramente expresados, que hacen las mismas discriminaciones que el hablante nativo. Es sabido que esta tarea requiere construcciones teóricas que van más allá de los preceptos *ad hoc* de nuestro conocimiento gramatical explícito. Probablemente, una situación similar ocurre en la teoría moral. No hay razón alguna para asumir que nuestro sentido de la justicia puede ser adecuadamente caracterizado por preceptos familiares de sentido común, o derivada de los más obvios principios de aprendizaje. Una correcta explicación de la capacidad moral, sin duda, implicará principios y construcciones teóricas que van mucho más allá de las normas y estándares citados en la vida cotidiana; esto puede, eventualmente, requerir también muchos cálculos sofisticados [\*Ver Noam Chomsky. *Aspectos de la teoría de la sintaxis*. Cambridge, Mass., M.L.T. Press. 1965, pp. 3-9] (Rawls, 2003, 41-42).

Rawls sugiere así una analogía entre el lenguaje y la moral, pues para él los principios fundamentales que nos permiten el desarrollo moral no son explícitos, así como no lo son, en Chomsky, los principios del desarrollo del lenguaje. Partiendo de aquí, los innatistas complementarán estas ideas con los trabajos actuales en psicología experimental y con los argumentos de la “carencia de estímulo” —llamados así, en tanto que se apoyan en situaciones morales en las que la experiencia no parece tener ningún tipo de participación; y, en consecuencia, tampoco las emociones—.

### 2.2.1 La competencia moral en Susan Dwyer

Susan Dwyer defiende que “[...] en un cierto nivel de abstracción, hay sorprendentes paralelismos entre el ejercicio y el desarrollo de la competencia moral, por un lado, y el ejercicio y el desarrollo de la competencia lingüística, por el otro” (Dwyer, 1999, 169). Ella se encamina así a defender el innatismo moral, retomando la estrategia lingüística y apoyándose en la postura de Rawls. Los lingüistas afirman que los niños se guían por reglas gramaticales que no podrían haber aprendido a través de la experiencia, porque los datos lingüísticos primarios a los cuales ellos están expuestos (las charlas de los adultos y las correcciones que los adultos hacen de sus charlas) no

proporcionan suficiente información para ayudarlos a seleccionar entre múltiples reglas posibles. Por analogía, Dwyer sostiene que los niños no reciben suficientes *inputs* morales primarios que expliquen su comprensión del dominio moral.

En particular, con estos *inputs*, ella se refiere a que los niños no reciben suficientes elementos de juicio de los adultos para distinguir entre reglas morales y reglas convencionales y que, sin embargo, lo hacen (Dwyer, 1999, 171). De hecho, Dwyer muestra la manera como la distinción moral/convencional ya está presente desde el tercer año de vida y no es explícitamente enseñada. Los padres no dicen a sus hijos cuáles reglas son morales y cuáles solo son convencionales, y los castigan por ambas. Los niños son castigados por golpear y morder y por violar reglas de etiqueta, pero de alguna manera ellos reconocen que las normas morales son diferentes. Es así como Dwyer sostiene que, en su infancia más temprana, los niños tratan las normas morales como más graves y menos dependientes de autoridades. Golpear y morder es para ellos erróneo sin importar que alguien se los diga, pero poner los codos sobre la mesa sería más bien correcto si se les permite. Todos los niños del mundo distinguen pues esta suerte de contraste sin que se les haya enseñado a hacerlo.

Por tanto, debido a esta distinción entre reglas morales y convencionales que los niños, sin una previa educación al respecto, parecen tener clara, Dwyer postula la existencia de una Gramática Moral Universal. La hipótesis de Dwyer dice, entonces, que: “[...] los seres humanos estamos equipados con una Gramática Moral Universal, un conjunto de principios abstractos, algunos de los cuales podrían ser sistematizados [*parametized*]” (Dwyer, 1999, 185). Partiendo de esta idea, la emoción o el razonamiento pueden jugar un papel en la emisión del juicio moral, pero nunca sin que la situación que suscita dicho juicio pase antes, como a través de una especie de filtro, por la Gramática Moral Universal.

En mi opinión, el argumento de Dwyer no es del todo demostrativo, puesto que se sostiene sobre una hipótesis no verificable, pero estoy de acuerdo con Prinz en que este argumento tiene una considerable plausibilidad *prima facie* (cf. Prinz 2007a, 267). Por esta razón, tal y como lo hace Prinz, es preciso considerar si esta plausibilidad hace que el argumento sea completamente sostenible. A este respecto, Prinz señala que aun cuando los padres no enseñen explícitamente la diferencia entre normas morales y convencionales, es de extrañar que aquí se haya dejado de lado la manera en la cual las reglas se transmiten implícitamente. Los padres tratan las transgresiones a las reglas morales de una manera más seria y las castigan más duramente. La infracción de las reglas morales es, a menudo, castigada con frases de poder e inducción empática, mientras que la infracción de las reglas convencionales es guiada con razonamiento y apelando a estándares de orden social. Además, la infracción de las reglas morales tiende a cobrarse con una carga emocional, porque a menudo involucra a una víctima. Así, por ejemplo, si la pequeña Susan muerde al pequeño Jesse, este llorará e, incluso, los padres de Jesse podrían enfadarse con Susan. Esta carga emocional implícita en estas situaciones es interiorizada por los niños, lo cual hace que juzguen que las reglas morales no dependen de una autoridad. Cuando ellos quieren morder, esto les puede causar algún tipo de malestar porque morder les produjo algún tipo de emoción negativa fuerte en un pasado reciente. Como resultado, los niños que han interiorizado emocionalmente las reglas morales dirán que no es correcto morder aun cuando alguna autoridad les diga que eso está bien. De acuerdo con esto, la idea de Prinz es que los niños están expuestos a suficientes correctivos implícitos para diferenciar entre reglas morales y convencionales —correctivos que no tiene en cuenta la postura de Dwyer— y, por tanto, no es necesario postular la existencia de una Gramática Moral Universal de carácter innato.

### 2.2.2 El órgano moral en Marc Hauser

Siguiendo la misma línea de Dwyer y apoyándose en los trabajos experimentales de John Mikhail (Mikhail, 2000, 2002), Marc Hauser (Hauser, 2006) ha sugerido recientemente que la respuesta emocional no es el mecanismo inicial mediante el cual producimos nuestros juicios morales, sino que es muy probable que poseamos una estructura aún más primigenia que se encargue de este propósito.

De acuerdo con Hauser, “[...] estamos dotados de una facultad moral que opera sobre las propiedades causales-intencionales de eventos y acciones, para conectarlos a consecuencias particulares” (Hauser, 2006, 215). Él y sus colaboradores postulan así la existencia de una facultad gramática moral universal, que consistiría de los principios y parámetros que son partes y lugares de dicha dotación biológica. La función de esta gramática, operada de una manera inconsciente, sería la de proporcionar un kit de herramientas para la construcción de posibles sistemas morales. A este respecto, Hauser afirma que:

El sistema moral particular resultante refleja detalles del entorno o cultura particular en la que se encuentre el sujeto, y detalles de un proceso de supresión del entorno mediante el cual parámetros particulares son seleccionados y combinados en las fases iniciales del desarrollo. Una vez que se establecen los parámetros que se combinan para lograr un sistema moral en particular, adquirir otro en una etapa posterior de la vida —llegando así a ser funcionalmente *bimoral*— es tan difícil y diferente como lo es aprender chino para un hablante nativo del inglés (Hauser, 2006, 15).

Así, debido a que esta facultad moral innata funciona de manera similar a como lo hace una gramática moral universal, Hauser defiende la misma propuesta de Dwyer con respecto a que se puede establecer una analogía entre esta facultad y la facultad para aprender un lenguaje —en términos de la gramática universal de Chomsky y apoyado en la analogía de Rawls—. De aquí que su argumento se conozca también como ‘la analogía del lenguaje’.

Hauser es consciente de la poca evidencia empírica que puede soportar esta analogía, pero también es consciente de la poca evidencia empírica que la podría refutar. Para apoyar su postura, Hauser et al., implementaron el ‘problema del tranvía’ —desarrollado por Mikhail (Mikhail, 2000 y 2002)—, con algunas variaciones, para aplicar un cuestionario en la web y recoger, así, miles de intuiciones morales de las personas. Este problema, en términos generales, es un experimento mental que, sustentado en los argumentos de la ‘carencia de estímulo’, busca determinar qué harían las personas y cómo justificarían su acción frente a la siguiente situación:

Mediante una de dos opciones, un transeúnte puede evitar que un tranvía que va directo hacia cinco trabajadores, acostados sobre los rieles, les quite la vida: *i.* puede activar una palanca que cambiará el carril del tranvía, pero esto hará que, inevitablemente, el tranvía le quite la vida a otro trabajador; *ii.* desde un puente peatonal, puede empujar a una persona que es lo suficientemente pesada como para detener el tranvía, pero esta persona, inevitablemente, también perderá la vida (cf. Mikhail, 2000).

En ambos casos, la decisión del transeúnte implicará quitarle la vida a una persona, pero le salvará la vida a cinco. De esta manera, indagando por cómo actuarían los navegantes de la web ante este tipo de situación, Hauser et al., encontraron que el 11% de las personas consideraron como correcto empujar al sujeto, mientras que el 89% consideró que lo correcto era cambiar el carril. A este respecto, Hauser señala:

Cuando se les preguntó a las personas sobre la justificación de sus decisiones, la gran mayoría fue incapaz de proveer una respuesta coherente, en particular porque el principio utilitarista de procurar el mayor bien es una constante en ambos casos, y los medios deontológicamente pertinentes implican el asesinato, que se presume como prohibido (Hauser, 2006, 218).

Hauser sugiere entonces que esta incapacidad de dar justificaciones coherentes puede tomarse como una evidencia de la existencia de principios morales innatos, principios que operan de manera inconsciente. Esto en razón de que, mediante un examen detallado de estos y otros casos de carencia de estímulo, se ha encontrado un patrón similar que emerge con algunos dilemas, mostrando una clara disociación entre

el juicio y la justificación, mientras que otros dilemas no mostraron disociación en lo absoluto. Así, según Hauser, en todos los casos, hay un principio operativo responsable de generar el juicio:

Por ejemplo, constantemente las personas juzgan los daños causados por intención como moralmente peores que los mismos daños causados por una acción conjeturada [*principio de intención*]; las personas juzgan los daños causados por una acción como peores que los mismos causados por la inacción [*principio de acción*] y, por último, se juzga los daños causados por contacto físico como peores que los mismos causados sin contacto físico [*principio de contacto*] (Hauser, 2006, 218).

En la justificación de estas distinciones, el equipo de Hauser encontró que la mayoría de las personas da una justificación necesaria para el principio de acción, un poco más de la mitad justifica el principio de contacto, pero muy pocos justifican el principio de intención. Estos resultados sugieren que el principio de acción y, en menor medida, el principio de contacto, no solo están presentes en la reflexión consciente, sino que parecen desempeñar un papel en el proceso que va de la percepción del evento al juicio moral y a la justificación moral. En cambio, el principio de intención, con su distinción entre consecuencias intencionadas y conjeturadas, parece ser inaccesible a la reflexión consciente. En consecuencia, cuando los dilemas morales involucran este último principio, las personas generan juicios morales intuitivos, usando procesos inconscientes para pasar de la percepción del evento al juicio moral. De acuerdo con esto, cuando las personas intenten justificar sus juicios, lo harán o de manera tal que no tengan una explicación coherente —dependiendo así de una especie de corazonada—, o proporcionando una explicación que es insuficiente —incompatible con afirmaciones anteriores—, o basadas en asunciones infundadas —que se harán en un intento por lidiar con su propia carencia de certeza—.

Para Prinz, siendo que las personas indagadas no habían oído hablar antes del problema del tranvía, sus respuestas podrían ser tomadas como una evidencia para el

innatismo. Pero el argumento de Hauser, de acuerdo con Prinz, funciona sobre el mismo principio de Dwyer: “Los autores identifican un elemento del juicio moral que parece estar muy generalizado, pero que no ha sido explícitamente enseñado” (Prinz, 2007a, 268) y de aquí concluyen que, entonces, debe tratarse de algo innato. Prinz sostiene, sin embargo, que este no es un argumento suficiente para el innatismo. Así, de un modo similar a como argumentó en contra de Dwyer, Prinz llama la atención sobre la evidencia de que se nos enseña tanto implícita como explícitamente a no hacer daño y a ayudar a otras personas. Siendo que claramente estas dos normas —evitar el daño y ayudar— entran en conflicto en los casos tranvía, la preguntas que surgen, de acuerdo con Prinz son: ¿cómo decidimos cuál norma aplicar? ¿Es esto determinado por principios innatos?

La respuesta, según Prinz, es que las intuiciones suscitadas por los casos tranvía son impulsadas por emociones, en tanto que si consideramos que es malo empujar a alguien hacia los rieles es porque esto suscita emociones intensamente negativas; además, si creemos que está bien cambiar el carril es debido a que la acción parece más benigna. Este análisis predice que las intuiciones tranvía fluctuarán, así, con la relevancia del daño y esta predicción parece estar corroborada por otros escenarios utilizados en los estudios del mismo Hauser. Así, por ejemplo, Prinz trae a colación un cuarto caso de Hauser, en el cual presentaron el escenario de la carrilera con una pequeña variación: el hombre pesado de la carrilera está de pie delante de un objeto pesado. A las personas se les dice que el objeto pesado servirá como barrera, impidiendo que el tranvía golpee a cinco personas, pero que el hombre pesado inevitablemente será asesinado en el proceso. Ahora el 72% de las personas consideró correcto accionar el interruptor; esto obedece, de acuerdo con Prinz, a que la atención se aleja de la víctima y recae sobre el objeto pesado, haciendo así que el daño se perciba como menor.

En líneas generales, el punto de Prinz es que todos los casos tranvía, presentados por Hauser et al., se pueden explicar a partir de las emociones y sin suponer, en consecuencia, que tenemos reglas innatas. En su opinión “La variación en nuestras intuiciones no refleja un conjunto de normas innatas precisas —una gramática moral— sino la facilidad con la que los sentimientos aprendidos pueden cambiar en función de la prominencia del daño” (Prinz, 2007a, 269). Además, Prinz señala que, pese a lo tentador que puede resultar, sería un error suponer que llegamos a posturas morales sin necesidad de mucha instrucción. Así, Prinz trae a colación los estudios de Hoffman (cf. Hoffman, 2000), para sustentar el no menospreciable hecho de la gran inversión de energía que hacen los padres educando moralmente a sus hijos. En estos estudios Hoffman (Hoffman, 2000, 141) calcula que, entre las edades de dos y diez años, los padres corrigen el comportamiento de sus hijos de cada seis a nueve minutos, lo cual equivale a cincuenta lecciones de conducta cada día. En este sentido, si la moralidad fuera innata, toda esta enseñanza debería ser innecesaria; pero es evidente, a juicio de Prinz, que esto no es así, pues la contribución de los padres es vital para la moral que el niño adquiere. En suma, Prinz considera que los argumentos innatistas no son nada concluyentes y que, pese a que sus consideraciones en contra —presentadas aquí— no son una prueba decisiva, la mejor manera de derrotar al innatismo es mostrando que el emocionalismo constituye una mejor explicación.

Ahora bien, yo pienso que una evidencia aún más fuerte que Hauser trae a su favor, y que constituye un argumento en contra de los emocionalistas, proviene de un experimento neurofisiológico. Con este, el equipo de Hauser apeló a las investigaciones realizadas por Koenigs et al., (Koenigs et al., en prensa) sobre el rol de las emociones en los juicios morales, y que involucran a pacientes con avanzados comienzos de daño ventromedial prefrontal [*Ventromedial prefrontal damage: VMPC*]. Este tipo de daño, que



ha sido cuidadosamente estudiado por Damasio, Tranel, Adolphs y Bechara (Bechara et al., 1994, 1997; Damasio, 1994, 2000, 2003; Tranel et al., 2000), indica un déficit en los pacientes para tomar decisiones tanto inmediatas como futuras. Una explicación de este déficit es que estos pacientes carecen del tipo de *input* emocional pertinente para la toma de decisiones, las cuales pueden experimentar personas sin este daño cerebral. Es decir, aquí se presupone que en las personas normales la toma de decisiones está íntimamente entrelazada con la experiencia emocional. En la ausencia del *input* emocional, la toma de decisiones carece de rumbo. A partir de aquí, el equipo de Hauser se pregunta, entonces: “¿cumplen las emociones un rol causalmente necesario en la generación de los juicios morales?” (Hauser, 2006, 217).

Para responder a este interrogante, utilizando al grupo de pacientes ya mencionado, a otro grupo con otros daños cerebrales y a un último sin daño alguno, el equipo de Hauser formuló un cuestionario con la forma general “¿Haría usted X?”. En un primer momento, se contrastaron dilemas no-morales con dos clases de dilemas morales: personales e impersonales. Los dilemas no-morales incluyeron situaciones en las que, por ejemplo, una acción de ahorro de tiempo podría ser potencialmente contrarrestada por un considerable costo financiero. Los dilemas personales e impersonales incluyeron casos que representaban opciones de dañar a una persona para salvar a muchas (el problema del tranvía). De esta manera, la característica distintiva fundamental entre personal e impersonal se basaba en que el primero requería algún tipo de daño, vía contacto físico, para lograr un bien individual o bienes individuales, mientras que el último no. El clásico problema del tranvía provee un caso simple. Así, en la versión impersonal, el transeúnte puede mover la palanca que hace que el tranvía se desvíe para que, en lugar de quitarle la vida a cinco personas, le quite la vida a una

sola persona; en la versión personal, el transeúnte puede empujar a una persona pesada delante del tranvía, quitándole la vida, pero salvando a cinco personas.

De acuerdo con Hauser, los tres grupos —los dos con daños cerebrales y el otro sin daños— mostraron los mismos patrones de respuesta para ambos dilemas no morales, así como para los dilemas morales impersonales. El caso que marcó la diferencia fue el del contexto de los dilemas morales personales: los pacientes VMPC fueron significativamente más propensos a decir que era permisible causar daño para salvar a un número mayor de otras personas, resultando una muy fuerte respuesta utilitarista —es decir que, con independencia de los medios, esto siempre es preferible a fin de maximizar el resultado global o la utilidad—.

Estos resultados sugieren que el déficit acumulado por los pacientes ventromediales no tiene un impacto global en los dilemas sociales y tampoco representa un impacto muy selectivo sobre los dilemas morales en general. Más bien, los daños en esta área impactan selectivamente los juicios acerca de los dilemas morales personales.

A partir de esto, Hauser afirma lo siguiente:

¿Qué aprendemos de este patrón de resultados, y especialmente del papel causal de las emociones? Dado que los dilemas morales impersonales son emocionalmente más destacados, podemos descartar la fuerte afirmación de que las emociones son causalmente necesarias para todos los dilemas morales. En lugar de ello, nos vemos obligados a concluir que las emociones juegan un papel más selectivo en una clase particular de dilemas morales, en concreto, aquellos relacionados con el daño personal (ibíd.).

En resultados separados, con respecto a las personas sin daño cerebral, todos los dilemas personales fueron clasificados como más emocionales que los dilemas impersonales. Pero Hauser y su equipo van todavía más lejos, pues examinando el rango de dilemas morales personales, revelaron una nueva distinción: algunos dilemas suscitan respuestas convergentes y rápidamente deliberadas (*bajo conflicto*), mientras que otros suscitaron respuestas muy divergentes y lentamente deliberadas (*alto*

*conflicto*). En consonancia, los pacientes ventromediales emitieron los mismos juicios morales que el resto de los grupos para los dilemas de bajo conflicto, pero juicios significativamente diferentes para los dilemas de alto conflicto. Una vez más, los pacientes ventromediales mostraron juicios muy utilitaristas cuando se contrastaron con los demás grupos. Así, por ejemplo, en un caso de bajo conflicto, en donde una adolescente quiere ahogar a su bebé recién nacido, todos los grupos estuvieron de acuerdo en que esto no sería admisible; en cambio, en un caso divergente como el de la decisión de Sophie (cf. Hauser, 2006, 218), en donde una madre debe permitir que uno de sus dos hijos sea probado experimentalmente o perdería a los dos, los pacientes VMPC declararon que el resultado utilitarista era permisible (es decir, permitir que se experimente con el niño), mientras que los otros dos grupos dijeron que no.

De acuerdo con esto, los casos de alto conflicto Vs. bajo conflicto parecen llamar la atención sobre una nueva distinción: entre situaciones de ayuda para sí mismo y para otros. Mientras el embarazo adolescente conlleva una decisión egoísta, la decisión de Sophie implica una consideración de daño a otros. Los pacientes VMPC mostraron el mismo patrón de juicios acerca de los casos de egoísmo como el grupo de los que tenían daños cerebrales, pero pusieron de manifiesto la respuesta utilitarista en la mayoría de los demás casos de ayuda a otros.

Para Hauser y su equipo de trabajo dos conclusiones surgen de esta serie de estudios. En primer lugar, el papel de las emociones en los juicios morales parece más bien selectivo, centrándose en lo que, de acuerdo con Hauser, podrían ser considerados los verdaderos dilemas morales: situaciones en las que no existen normas sociales claras para juzgar lo que es moralmente correcto o incorrecto y cuando el contexto es intensamente emocional (cf. *ibíd.*). Una interpretación de este resultado es que, a falta de regulación emocional normal, las personas VMPC fallan al experimentar el conflicto

clásico entre el cálculo que permite un análisis utilitarista o consecuencialista y el sistema que se dirige por reglas o principios deontológicos o no consecuencialistas (Greene y Haidt, 2002; Greene et al., 2004; Hauser, 2006). De este modo, Hauser afirma que cuando los *inputs* emocionales se ausentan, el razonamiento consecuencial aflora, como si los sujetos fueran ciegos a las normas deontológicas y no consecuencialistas. En segundo lugar, Hauser concluye que de esta evidencia se sigue el rechazo de una versión fuerte de la criatura humeana tanto como de una versión fuerte de la criatura rawlsiana. Esto en razón de que las emociones ni son causalmente necesarias para generar todos los juicios morales, ni son irrelevantes para generar todos los juicios morales. Por el contrario, en algunos de los dilemas morales, tales como aquellos que caen bajo la categoría de impersonales, como personal/bajo conflicto/egoísmo, las emociones parecen jugar poco o ningún papel. En contraste, para dilemas personales/alto conflicto/ayuda a otros, las emociones parecen jugar un papel causal determinante.

Dicho esto, es de señalar entonces que si bien la postura de Hauser rechaza la idea de que una estructura emocional ocupe el primer lugar en la producción del juicio moral, su innatismo no constituye una defensa radical de la postura contraria.

Sin embargo, Hauser señala que esta última conclusión, según la cual las emociones parecen jugar un papel causal determinante, debe ser examinada a la luz de nuestra más bien limitada comprensión de la forma y el contenido representacional de las emociones, así como de sus bases neuronales. Esto porque, como lo señalamos anteriormente, la evidencia se apoya aquí enteramente en la afirmación de que la corteza prefrontal ventromedial es responsable del tráfico de las experiencias emocionales a los procesos de toma de decisión. Pero si resultara que otros circuitos neuronales están críticamente involucrados en el procesamiento emocional, y que ellos

están intactos, entonces el patrón relativamente normal de respuestas en los casos impersonales, así como en los casos personal/bajo conflicto/auto ayuda, sería de esperarse (cf. Hauser, 2006, 217-218).

Ahora bien, si bien es cierto que la estructura en cuestión u órgano moral, según Hauser, no es emocional, tampoco —en mi opinión y de acuerdo con la exposición que hemos visto del autor— sería racional, propiamente hablando, puesto que su teoría postula una capacidad más bien innata e inconsciente que actuaría después de la percepción de un evento moralmente relevante y justo antes de que la emoción y el razonamiento moral tengan lugar. Por esta razón, si por racionalidad entendemos el procesamiento consciente de información, entonces no sería del todo correcto subsumir a Hauser bajo el rótulo de “racionalista moral”. Esto podría resultar irrelevante, pero examinándolo con mayor atención, podría brindar un pequeño eslabón para sugerir dos cosas. Por un lado, esto podría sugerir que la postura de Hauser no se ubica en un debate entre emocionalistas y racionalistas —y que, en consecuencia, tal debate no existe, siendo él el ‘supuesto’ mayor exponente de la postura racionalista— sino que, por otro lado, se trata más bien de una indagación acerca del proceso, aún más primigenio, que permite que se activen las estructuras emocionales y racionales que, a su vez, dan lugar al juicio moral. En segundo lugar, y en apoyo del punto anterior, es de señalar una cierta disonancia que hay entre la manera que Hauser presenta su teoría y el contenido de la misma, puesto que se trata de postular un órgano (físico) con códigos racionales cuya estructura no se puede determinar. Esto en razón de que la idea fundamental de Hauser es equiparar la estructura que postula con un órgano del cuerpo, y de aquí el título de su artículo: “El hígado y el órgano moral”. Él cree —sin que, en su opinión, lo tenga muy claro— que si el órgano que postula ofrece una herramienta universal para construir sistemas

morales particulares, entonces, en contra de la intuición cotidiana, uno debería estar dispuesto a que le trasplanten partes del cerebro —de manera similar a como se trasplanta un hígado—, en el caso de que tenga estas partes dañadas; tanto un hígado sano como un cerebro sano cumplirían muy bien su trabajo (cf. Hauser, 2006, 220).

## 2.3 EL HÍBRIDO EMOCIONALISMO-RACIONALISMO MORAL

Joshua Greene es un filósofo de la Universidad de Harvard que, mediante experimentos conductuales, neuro-imagen funcional (fMRI) y otros métodos neuro-científicos, también ha investigado los procesos que dan lugar al juicio moral y a la toma de decisiones. La evidencia, para Greene y sus colaboradores, nos muestra que el propósito general de la investigación en este campo debe ser doble. Por un lado, se trata de entender la manera en que los juicios morales son formados por procesos automáticos —tales como “reacciones viscerales” emocionales— y, por otro lado, entender cómo estos mismos juicios son controlados por procesos cognitivos —tales como el razonamiento y el auto-control— (Greene, en prensa, *Dual-process*). De aquí que la afirmación principal de Greene et al., consista, entonces, en que tanto la emoción como la razón desempeñan un papel fundamental en el juicio moral.

En consecuencia, las indagaciones de Greene et al., se extienden más allá de la pregunta por cuál es, entre la razón o la emoción, la causa fundamental de los juicios morales, a preguntas no menos pertinentes, tales como: ¿cuáles son los correlatos neuronales de la razón y la emoción?, ¿cuál es la naturaleza de su interacción? y, por supuesto, ¿cuáles son los factores que modulan sus respectivas influencias conductuales en el contexto del juicio moral? (cf. Greene et al., 2001, 2105).

### 2.3.1 El proceso dual de la moral en Joshua Greene

Para responder a dichas preguntas, Greene et al. (2001) hicieron dos estudios de imagen de resonancia magnética funcional (fMRI), utilizando dilemas morales como guías; dilemas que consideraban casos, como los que ya hemos visto, acerca del problema del tranvía. En estos estudios, las personas examinadas mostraron una activación significativa en las áreas emocionales del cerebro —Giro frontal medio [*Medial frontal gyrus*], Giro cingulado posterior [*Posterior cingulate gyrus*] y Giro angular [*Angular gyrus*] tanto derecho como izquierdo— cuando se les preguntó si era apropiado empujar a alguien delante del tranvía para salvar a cinco trabajadores. Sin embargo, estas personas mostraron activaciones emocionales mucho más bajas cuando se les preguntó si era apropiado cambiar el carril para que, en lugar de matar a cinco trabajadores, el tranvía matara a uno nada más. En este último escenario, las áreas del cerebro que más presentaron activación fueron aquellas asociadas con el procesamiento de la memoria —Giro frontal medio [*Middle frontal gyrus*] derecho y Lóbulo parietal [*Parietal Lobe*] tanto izquierdo como derecho—. De acuerdo con estos resultados, los autores sugieren que el razonamiento moral está dirigido por dos procesos disociables: un proceso emocional y un proceso racional.

Ahora bien, siendo que este contraste en las activaciones cerebrales tuvo lugar, a su vez, a raíz de la diferencia de los dilemas morales considerados, Greene et al., señalan —de una manera muy similar a como lo hizo Hauser— que en el juicio moral los eventos o situaciones moralmente relevantes que percibe el sujeto también son un factor determinante —esto es, que el tipo de *input* determinará, a su vez, el tipo de estructura, emocional o racional, que entrará en juego para emitir el juicio—. Con respecto a esto, ellos dividieron los dilemas morales (los *inputs*) en dos grupos: los dilemas personales y los dilemas impersonales. En los primeros se ubica la situación en

la que el sujeto indagado debe empujar a una persona para salvar la vida de los cinco trabajadores; en el segundo grupo, por su parte, se ubica la situación en la que la opción es cambiar el carril para que una sola persona pierda la vida y no cinco. De esta manera, los dilemas personales quedaron vinculados a las áreas emocionales —puesto que ante ellas el papel de las áreas del razonamiento fue muy poco e, incluso, nulo—, mientras que los dilemas impersonales se vincularon con las áreas del razonamiento —entendidas, en este caso, como las áreas encargadas del procesamiento de la memoria—.

Sin ir muy lejos, el equipo de Greene defiende así que la clase de dilemas morales que se percibe hace que varíe, a su vez y sistemáticamente, el tipo de proceso involucrado en el juicio moral, bien sea un procesamiento emocional o uno de tipo racional. Prinz, por su parte, interpreta los datos que apoyan esta conclusión de una manera diferente: las emociones determinan los juicios morales en ambos casos.

La postura de Prinz se sostiene ahora en que en los casos del tranvía seguimos una regla tácita, respaldada por una emoción, que dice que matar es malo, y una regla un poco más débil, pero igualmente respaldada por una emoción, que dice que salvar vidas es bueno. Bajo este planteamiento, en el caso de ‘empujar’ al hombre pesado, podemos imaginar la muerte de una manera muy vívida y recibimos así el fuerte golpe emocional de la regla “¡No matar!”, que supera las emociones más débiles asociadas con la regla “¡Salvar la vida!”. Por su parte, en el caso de accionar la palanca, no nos imaginamos muy vívidamente el daño que estamos causando, razón por la cual la regla “¡Salvar la vida!” puede guiar nuestras acciones. Para Prinz, aquí los números son de vital importancia, puesto que nosotros calculamos que accionar la palanca se traducirá en más vidas salvadas y esto deriva, así, en una preferencia emocional por accionar la palanca. Las activaciones en la memoria de trabajo resultan, entonces y para Prinz, del



hecho de que nuestra decisión depende de pensar en los números. Así, dice Prinz, “[...] fríamente calculamos cuál curso de acción puede salvar más vidas, pero una vez sabemos que es moralmente mejor accionar la palanca, *este* juicio puede ser respaldado por una respuesta emocional” (Prinz, 2007a, 24). Prinz considera que esta explicación es consistente con los datos obtenidos por Greene et al., pues recuerda que en dichos datos las emociones estuvieron activas en ambos casos —en el de empujar y accionar la palanca—. De esta manera, las emociones resultan más intensas en el caso de empujar, porque empujar a alguien a su muerte es una actividad demasiado evocadora.

El argumento principal de Prinz es, entonces, que deliberamos sobre los dilemas morales oponiendo emociones a emociones. Esto significa que algunas normas entran en conflicto porque tienen una fuerza emocional diferente, y en esta contienda las emociones más fuertes resultan siempre ganadoras. Es así como Prinz predice que los juicios cambiarían notablemente si cambiáramos los escenarios de los casos tranvía en formas emocionalmente significativas. He aquí la predicción:

Quando las personas dicen que es moralmente aceptable accionar la palanca para salvar a cinco personas y matar a una, ellas se imaginan que la palanca a ser accionada está lejos de los carriles. Ahora, supongamos que les decimos a estas personas que la palanca se encuentra a pocos centímetros de la persona que matarían si la accionan. Imagínese usted mismo esa situación. Un hombre está atado al carril a su lado. Usted no puede liberarlo. Él se retuerce alrededor y grita de terror. Usted sabe que hay cinco personas en otro carril, que están a cierta distancia, y usted sabe que el tranvía se dirige hacia ellas. ¿Sacrificaría usted a la persona a sus pies? ¿Podría ser esto moralmente aceptable? Aquí, creo que las intuiciones cambiarían. Esto es más parecido al caso de empujar. Las personas que no habían estado expuestas a muchos de estos ejemplos podrían, por defecto, tener un serio recelo moral acerca de sacrificar la vida de alguien que tienen a pulgadas de distancia. Las emociones fuertes suscitadas por la proximidad de la víctima, predigo yo, influirán en el juicio (Prinz, 2007a, 25).

Para la plausibilidad de esta predicción, Prinz encuentra un apoyo en los resultados de otra variante registrada por Greene. En esta, nos encontramos ahora en una sala de control y nos percatamos de que un tranvía se dirige hacia cinco personas;

podemos evitar que esto pase accionando una palanca que abrirá una compuerta para que una persona, parada en el puente peatonal, caiga en el camino del tranvía y lo descarrile —es de anotar que, en este escenario, no se requiere ningún contacto físico con la víctima—. Greene et al., (de próxima aparición), encontraron que la mayoría de la gente piensa que es lícito matar al hombre en el puente; así, si a las personas se les decía que tenían que empujar al hombre desde el puente, solo el 31% decía que era permisible, mientras que si ahora se les dice que solo hay que accionar una palanca que abre una compuerta, el 63% piensa que es admisible. De aquí, según Prinz, se sigue que no somos esclavos de un principio de que matar es peor que dejar morir, pues aunque normalmente adherimos a este principio, un cambio en la intensidad emocional puede llevarnos a apoyar claras violaciones del mismo. Es así como, sin mayores remordimientos, podemos dejar que alguien muera si no hay ningún tipo de contacto con nosotros —como ocurre, tan a menudo, con las crisis y hambrunas distantes en todo el mundo—. Por tanto, de acuerdo con Prinz, “la disminución de la intensidad emocional del método de matar duplica el índice de aprobación” (ibíd.).

## 2.4 ¿UN DIÁLOGO DE SORDOS?

Tal y como hemos visto hasta aquí, el uso de métodos empíricos para explorar las preguntas tradicionales en la teoría moral todavía está en su etapa de gestación y queda mucho por aprender. En este capítulo he examinado tan solo una pequeña muestra del trabajo general en este ámbito. Sin embargo, creo que esto ha servido para ilustrar la manera en que este trabajo puede arrojar luces importantes sobre áreas de investigación que tradicionalmente han sido de interés filosófico.

En mi opinión, la solución al debate que en este ámbito pretenden librar las posturas racionalistas y emocionalistas no parece del todo excluyente. Quiero llamar la

atención sobre dos aspectos que encuentro de vital importancia, en tanto que sugieren una manera de dirimir el conflicto. En primer lugar, cabe resaltar que, en este debate, las posturas surgen de la elección de diferentes situaciones prototípicas: los defensores del enfoque racional-innato se centran en los dilemas morales sofisticados, esto es, sobre aquellos en los que entran en conflicto posturas generalmente aceptadas y, en consecuencia, hacen más difícil tomar una decisión (cf. ‘dilemas impersonales’ en Hauser y Greene); por su parte, los defensores del enfoque intuitivo-emocional parecen centrarse en las reacciones que, casi que de manera inmediata, podemos tener frente a las infracciones morales de otras personas e, incluso, de nosotros mismos (cf. ‘dilemas personales’ en Hauser y Greene). Mi punto es, entonces, que la elección de uno u otro dilema como la situación moral prototípica parece arbitraria, y es claro que es ella quien tiene un impacto significativo en el juicio moral resultante.

En segundo lugar y más importante aún, es central el hecho de que unos y otros defiendan los resultados de acuerdo con su postura. Este aspecto opaca, así, al anterior, puesto que si bien es cierto que la elección de la situación moral prototípica parece arbitraria, no se aportaría mayor cosa utilizando una misma situación siendo que las dos posturas interpretarán los resultados a su acomodo —y remitiéndose a agregar, con un extraño tono de modestia, “hay mucho por investigar”—.

Así, por ejemplo, volvamos por un momento a la respuesta de Prinz a Greene. *Prima facie*, parece muy cierto que la disminución de la intensidad emocional del método de hacer daño aumenta el índice de aprobación de dicho daño. Pero, *prima facie*, no parece menos cierto que un cálculo que arroje como resultado una disminución en la intensidad del daño mismo aumentará el índice de aprobación de dicho daño. ¿De qué le sirve al debate que el emocionalista contemple factores tales como la cantidad del daño si los termina interpretando, de una u otra forma, en función

de las emociones? Siendo que no le sirve de nada, el emocionalista se niega a escuchar los planteamientos racionalistas, haciendo que este debate no resulte ‘productivo’.

En consecuencia, considero que el mejor argumento en este debate es el que busca integrar los dos modelos (junto con la riqueza de la evidencia empírica que cada uno ha generado); aquel que contempla los *inputs* o contextos en los que cada uno tiene mejor aplicabilidad. De esta manera, pienso que promete mucho más el modelo de Greene, en el que interactúan la razón y la emoción —de acuerdo con las diferentes circunstancias morales que percibe el sujeto—. Además, considero que es de vital importancia tener en cuenta que, en ambos lados de la discusión, emocionalistas y racionalistas —como Haidt y Hauser, en particular— han aceptado la participación de ambos componentes en los juicios morales. Así, parece que sólo Prinz es quien ha decidido defender su postura de una manera radical y unilateral —desde el lado emocionalista—, tan solo por encontrar que se trata del argumento a la mejor explicación. Yo encuentro, así, que un modelo que complemente los procesos racionales con los emocionales puede proporcionarnos una comprensión más enriquecedora del proceso del juicio moral, puesto que contempla la complejidad de las situaciones morales encontradas en la cotidianidad y sugiere que, ante dicha complejidad, nuestras respuestas morales no deben ser menos complejas.

---

### 3. LA TESIS EMOCIONALISTA EN LA EXPLICACIÓN COGNITIVA DEL JUICIO DE GUSTO ESTÉTICO

En el segundo capítulo espero haber mostrado lo inapropiado que resulta preguntarse si el proceso mental a partir del cual expresamos valoraciones de tipo moral depende de las emociones o de principios racionales, de manera excluyente, pues ambos parecen jugar un papel igualmente importante en este proceso. En este capítulo, mi idea es socavar el carácter unilateral del emocionalismo desde otra área de la teoría de los valores, igualmente relevante para esta discusión, tal y como es la valoración estética. Pienso que es justamente este el ámbito desde el que quizás más podemos defender esta intuición, puesto que cuando hablamos de la apreciación de una obra de arte el conocimiento juega un papel fundamental. Es pues este el propósito del presente capítulo, la refutación de la segunda parte de la tesis emocionalista de Prinz.

Tal y como lo señalé anteriormente, encuentro que la tesis emocionalista, en este segundo componente, malentende la manera en que funcionan las obras de arte y sugiere una comprensión sesgada de las mismas, al ubicar en un segundo plano e, incluso, obviar el papel que cumple el conocimiento en la apreciación de una obra de arte en cuanto obra. Por esta razón, mi idea principal aquí es defender que las obras de arte funcionan, en esencia, como una forma de ‘expresión’, de ‘expresión metafórica’,

en particular. De aquí que una respuesta emocional no sea suficiente para apreciar estéticamente una obra de arte; la apreciación misma requiere, además, del entendimiento de la manera en que funciona la expresión metafórica, esto es, entender el sentido en que la obra de arte funciona como obra de arte.

Así pues, a manera de conclusión, gran parte de la argumentación subsiguiente está encaminada a mostrar en qué sentido podemos afirmar que, además de valorar las obras de arte como buenas o malas, grandiosas u horrorosas, también las valoramos como correctas e incorrectas, dando así una explicación cognitiva del juicio de gusto estético que involucra criterios de corrección. De acuerdo con esto, intento fijar así la atención del lector sobre la poca solidez de los límites que se han establecido entre el arte y la ciencia, sin pretender llegar a conclusiones contundentes y, más bien, sugiriendo algunos derroteros a seguir en los ámbitos de la apreciación del arte y la educación artística y el museo, y la relación entre estos y el trabajo científico.

Para este efecto, luego de presentar la tesis emocionalista de Prinz, expondré dos argumentos principales, uno de tipo técnico y otro de tipo más bien intuitivo. El primer argumento, el de tipo técnico, está encaminado principalmente a mostrar el sentido en que podemos hablar de criterios de corrección en las artes. Para esto, me apoyaré en los aportes que Nelson Goodman —en su libro *Los lenguajes del arte*— y Catherine Z. Elgin —en *With Reference to Reference*— han hecho sobre el problema que subyace a nuestro entendimiento y subsiguiente apreciación de las obras de arte, esto es, el problema de la referencia en las artes. Con el segundo argumento, por su parte, se trata de señalar de una manera muy breve las indeseables consecuencias que, en la práctica, conlleva la defensa de que el arte es una especie de vehículo encaminado a la conmoción de las emociones o, como variante de esta postura, que la apreciación estética consiste en una respuesta preponderantemente emocional —tal y como defiende Prinz—.

### 3.1 ARTE Y EMOCIÓN

En su artículo “Emotion and Aesthetic Value”, Jesse Prinz argumenta a favor de una explicación emocional de la valoración estética (cf. 2007b). Para él, tanto la ética como la estética son de un dominio normativo, y este dominio, como vimos anteriormente (cf. *infra.*, cap. 1), concierne a la forma en la que el mundo debería ser, mas no a la forma en la que es. Recordemos que Prinz, apoyándose en las tesis de Hume, defiende que juzgar una acción como moralmente mala es el resultado de “experimentar desaprobación” y que juzgarla como buena es el resultado de “experimentar aprobación”. Este tipo de juicios, en Prinz y en la tradición de la filosofía moral británica es, pues, la expresión de una emoción ante una acción moralmente relevante; y de aquí que, bajo esta postura, la moral se construya emocionalmente. En el artículo en cuestión, Prinz defiende que una explicación emocional de la valoración estética es igualmente prometedora. Según él, aunque “[...] hay diferencias importantes entre los dos ámbitos, ambos tienen un fundamento afectivo” (Prinz, 2007b, 1); porque de un modo similar a como valoramos una acción moralmente relevante, también valoramos las obras de arte como buenas o malas, como grandiosas u horrorosas. El autor se referirá, así, a una valoración positiva de una obra de arte como una apreciación estética de esta obra, y a una valoración negativa como una depreciación estética.

Prinz está interesado, entonces, en la clase de estado mental en la que consiste la apreciación, y su indagación es, de esta manera, una sobre psicología estética. De aquí que él se pregunte “[...] cuando elogiamos una obra de arte, cuando decimos que tiene un valor estético, ¿en qué consiste nuestro elogio?” (*ibíd.*); y, sin ir muy lejos, su respuesta es la siguiente: “No afirmaré que las obras de arte expresan emociones o, incluso, que ellas necesariamente provocan emociones. Únicamente afirmaré que

cuando apreciamos una obra la apreciación consiste en una respuesta emocional” (ibíd.).

Para la defensa de esta tesis, Prinz divide su artículo en dos partes: en la primera se ocupa de argumentar en qué sentido las emociones están involucradas en la apreciación estética, y luego considera, en particular, cuáles emociones están involucradas. A continuación, haré una breve sinopsis de los argumentos de la primera parte del artículo y, tan solo de una manera muy general, expondré las conclusiones a las que Prinz llega en la segunda parte. Esto en razón de que mi interés aquí es entender qué es aquello que quiere decir Prinz al afirmar que la apreciación consiste en una respuesta emocional; propósito que bien se puede lograr examinando la primera parte de su texto. De acuerdo con esto, mi idea es responder, principalmente, a la siguiente pregunta: ¿Quiere decir Prinz que la apreciación estética es el resultado de una respuesta emocional o que dicha respuesta es suficiente para la apreciación estética, esto es, para que se profiera un juicio estético? Pienso que es importante dar respuesta a este interrogante, puesto que de ser lo primero, la propuesta de Prinz bien podría contemplar una participación igualmente importante del conocimiento en la apreciación estética. Sin embargo y como defenderé aquí, la idea del autor es defender lo segundo, esto es, que la respuesta emocional es suficiente para que surja la valoración estética; tesis que, en mi opinión, malentiende el funcionamiento de las obras de arte y resulta unilateralmente reduccionista.

### **3.1.1 Una teoría afectiva de la apreciación estética**

Prinz comienza ofreciendo algunas razones para pensar que la apreciación es un estado emocional. Es de anotar que, para él, no hay ningún argumento contundente para esta conclusión. “Más bien —según nos dice—, uno puede defenderlo como un argumento



a la mejor explicación, puesto que la hipótesis de que la apreciación tiene un fundamento afectivo sistematiza un buen número de observaciones que, de otra manera, sería muy difícil explicar” (Prinz, 2007b, 2).

Dichas razones, acerca del enlace entre la apreciación estética y la emoción, Prinz las divide en varias categorías, que en aras de la brevedad bien podríamos sintetizar de la siguiente manera:

a. *Argumento de la introspección*: Este argumento señala que las emociones ocurren simultáneamente con la apreciación, puesto que, mediante una simple introspección, podemos inferir que cuando vemos obras de arte y llegamos a una valoración vemos que es perfectamente obvio que estamos teniendo una reacción emocional: “El buen arte puede ser excitante, el mal arte puede ser deprimente” (ibíd.). Sin embargo, este argumento no es nada concluyente —cosa de la que el mismo Prinz es consciente—, puesto que las experiencias introspectivas de una persona difieren de las de otra en grado sumo.

b. *Argumento de la neuroimagen*: Estudios de imagen de resonancia magnética funcional (fMRI) —cf. Kawabata y Zeki (2003), Vartanian y Goel (2003), y Cela-Conde et al. (2004)— señalan que en la contemplación de obras de arte se activan: la corteza orbitofrontal, la circunvolución cingulada anterior e izquierda, y el polo temporal, entre otras, que son áreas asociadas con la emoción. Así, si bien es sugestiva esta clase de evidencia, a la lista también habría que añadir —como no lo hace Prinz— el lóbulo frontal, que está asociado con la memoria y el conocimiento y cuya activación está igualmente presente en la apreciación (cf. Greene et al., 2004).

c. *Argumento de la influencia del estado emocional en la apreciación*: Test y experimentos psicológicos —cf. Mealey y Feis (1995), y White et al. (1981)— muestran que el

estado emocional de los espectadores influye en sus preferencias estéticas. En mi opinión, este tipo de influencia solo ocurre cuando el estado emocional en cuestión es demasiado intenso y, por lo general, la preferencia no es perdurable; además, Prinz acepta que estos experimentos —grabaciones de audio emocionalmente evocadoras para incrementar ciertas valoraciones de atracción física— no fueron aplicados a la apreciación de obras de arte (cf. Prinz, 2007b, 2).

d. *Argumento de la influencia de la personalidad en la apreciación:* Test psicológicos —cf. Furnham y Walker (2001), y Rosenbloom (2006)— muestran que personas con rasgos de personalidad diferentes tienen gustos diferentes. A partir de este hecho —que en verdad no parece requerir ningún tipo de evidencia empírica—, el argumento de Prinz descansa en la poco fundamentada conclusión de que los rasgos de personalidad pueden ser interpretados como disposiciones emocionales y que, por consiguiente, estos hallazgos pueden señalar un enlace entre la emoción y la preferencia (cf. Prinz, 2007b, 2-3).

e. *Argumento del condicionamiento de la apreciación:* Este argumento descansa en el bien conocido hecho de que exposiciones repetidas a un estímulo inducen afectos positivos (cf. Prinz, 2007b, 3). Sin embargo, si bien parece cierto que la familiaridad —incluso sin recuerdo— induce afectos positivos, y que los afectos positivos incrementan la preferencia estética, esto no explicaría, entonces, el igualmente bien conocido hecho de que podemos mostrar una alta preferencia por una obra que nunca habíamos visto.

f. *Argumento de la reducción del interés estético debido a la reducción de las emociones:* Existe evidencia —cf. Chapman y Chapman (1983); y Bagby et al., (1994)— de que si disminuyen las emociones —problemas con la anhedonia y la alexitimia—, hay una reducción correspondiente en el interés estético; las personas que carecen de

emociones positivas fuertes tienden a tener menos apreciación por las experiencias estéticas que otros (cf. Prinz, 2007b, 3). Sin embargo, pienso yo, estas personas, por un lado, no representan casos que podríamos denominar ‘normales’ y, por otro lado, bien podría ser que ellas desarrollen gustos estéticos diferentes a los de la mayoría de las personas.

g. *Argumento de la estética popular sobre la subjetividad del gusto*: Este argumento apunta a señalar el hecho fehaciente de la variabilidad en el gusto, el cual —desde la estética popular— se explica diciendo que el valor estético está en la mirada del espectador. Prinz está de acuerdo tanto con el subjetivismo como con el relativismo que encierra esta creencia, y de aquí infiere el carácter contextual del gusto estético. Esto se puede explicar, en su opinión, porque “[...] las preferencias estéticas están basadas en emociones, y las emociones pueden ser condicionadas de manera diferente en lugares culturales diferentes” (cf. Prinz, 2007b, 3-4). Pero no son solamente las emociones sino también las creencias las que pueden, de manera contextual, configurar los diferentes gustos.

h. *Argumento de la estética popular sobre la objetividad del gusto*: En contraste con el argumento anterior, la estética popular también tiende a ser objetivista con respecto al arte: “Tendemos a pensar que las obras de arte podrían ser bellas incluso si nadie continuara admirándolas (Nichols, \*\*\*), y también tendemos a pensar que algunas personas tienen mejor gusto que otras” (cf. Prinz, 2007b, 4). Prinz considera que esta forma de objetivismo estético también se puede usar para equiparar la apreciación con la emoción:

Esto en razón de que tendemos a proyectar nuestras emociones en el mundo. Supongamos que una pintura nos hace sentir bien y, entonces, nos preguntamos si la pintura podría seguir siendo bella aun si las personas no reaccionaran ante ella. Cuando imaginamos el caso, continuamos imaginando la pintura y, mientras imaginamos la pintura, seguimos sintiéndonos bien. Esto nos lleva a pensar que la pintura es intrínsecamente buena. Y si las obras de arte

pueden ser intrínsecamente buenas, entonces puede haber hechos estéticos objetivos. Así, irónicamente, la misma cosa que hace al gusto subjetivo y relativo, también nos engaña al hacernos pensar que el valor estético es objetivo (cf. *ibíd.*).

Sin embargo, no es esta la única explicación posible para el objetivismo. Así, por ejemplo, y trayendo a colación los argumentos de Hauser expuestos en el capítulo anterior, bien podríamos decir que es esta una evidencia de la existencia de principios universales racionales e innatos —que operan de manera inconsciente en la apreciación— (cf. *infra.*, secc. 2.2.2).

i. *Argumento de la debilidad de la tesis contraria:* Si defendemos que la apreciación no es afectiva, entonces, cabría preguntarse: ¿qué es? La respuesta más obvia podría ser que se trata de un proceso racional —o el producto de un proceso racional—. Ante esta respuesta, Prinz señala que ninguna cosa que podamos saber acerca de una pintura (datos acerca de su génesis, forma o contenido, por ejemplo) parece ser suficiente para determinar que la obra es buena. Esto porque “[...] se puede saber que la obra goza de un buen equilibrio compositivo, es original y está hábilmente ejecutada, pero no se puede inferir que la obra es buena sobre estas bases a menos que se valoren el equilibrio, la originalidad y la habilidad. El valor de estas cosas no puede ser un hecho descriptivo adicional sobre ellas, porque para cualquier hecho descriptivo puede haber una pregunta acerca de si es digno de apreciación” (*ibíd.*). Si este es el caso, yo creo que el argumento también puede correr en sentido contrario: una obra de arte podría despertar en nosotros emociones muy fuertes, pero no por ello podemos inferir que la obra sea buena sobre estas bases a menos que valoremos estas emociones.

Dicho esto, debo decir que estoy de acuerdo con Prinz en que ninguno de sus argumentos es concluyente para defender que la apreciación o valoración estética es un estado emocional. A lo sumo, algunos de estos argumentos muestran la participación de las emociones en dicha apreciación, pero no que en ella esta participación sea

necesaria y mucho menos suficiente. Creo que es hora de preguntar, entonces, qué es lo que Prinz quiere defender, exactamente, al afirmar que la apreciación estética tiene un fundamento emocional: ¿cómo ocurre este proceso?, ¿qué papel cumplen las emociones en él? y ¿qué es, *stricto sensu*, aquello que se entiende aquí como apreciación estética?

### 3.1.1.1 *El modelo emocional del proceso de la apreciación estética*

Prinz propone un proceso de la apreciación estética en el cual distingue dos fases: una respuesta inicial ante la obra y luego una valoración de la obra —que se nutre de dicha respuesta—. Para él, ambas fases implican emociones (cf. Prinz, 2007b, 5).

Así, de acuerdo con el autor, la fase de respuesta es aquella en la cual percibimos la obra y reaccionamos ante sus características. En muchos casos la reacción está impulsada por factores perceptuales de los que somos totalmente inconscientes; pero el punto principal es que estos factores —consciente o inconscientemente— pueden suscitar emociones en nosotros, porque nos recuerdan cosas emocionalmente significantes en el mundo real. Aquí también entran en juego prejuicios implícitos que podemos tener hacia determinadas características composicionales.

Ahora bien, Prinz acepta que “esta fase de respuesta también puede ser afectada de arriba-abajo por el conocimiento” (ibíd.). Esto en razón de que si sabemos que una imagen fue producida de una determinada manera, esto podría excitarnos más; y, así, las creencias también pueden afectar la atención y la interpretación. Nótese que esta afirmación puede ser bastante desconcertante, en tanto que podría condenar ya la tesis de Prinz, pero permítaseme examinar primero la segunda fase del proceso y en breve volveré sobre este punto.

La segunda fase de la apreciación estética, de acuerdo con Prinz, es la valoración, propiamente hablando; fase en la que consideramos las respuestas de la primera fase a la luz de nuestros valores estéticos. Para él, “[...] un valor estético es una regla guardada a largo plazo en la memoria, que puede ser esquematizada de esta manera: si una obra W tiene una característica F, entonces, en esta medida W es buena al grado N” (ibíd.). Así, por ejemplo, si una obra cuenta con destreza técnica, entonces la valoraremos de acuerdo con el grado de destreza que percibamos en ella.

Prinz reconoce que en esta fase también aportamos más conocimiento de trasfondo, en tanto que el grado de valoración puede responder a preguntas tales como: ¿La obra es original? ¿Responde de manera interesante a otras obras en la historia del arte? (ibíd.). Y, sin embargo, él señala dos objeciones a esta participación del conocimiento. En primer lugar, Prinz considera que estas formas tan explícitas de deliberación podrían ser muy raras, puesto que la investigación muestra que cuando razonamos explícitamente sobre nuestras preferencias, hacemos malas elecciones que llegamos a lamentar (cf. Wilson et al., 2003). En segundo lugar, Prinz trae a colación la evidencia de que el razonamiento explícito aquí es *post-hoc*, ya que las personas esgrimen explicaciones de por qué prefieren una de dos imágenes incluso cuando los experimentadores intercambian las dos imágenes en secreto, de modo tal que las personas terminan generando razones para preferir una imagen que no fue la que ellos, de hecho y minutos antes, escogieron como preferible. Esto sugiere, así, que la valoración involucra, con frecuencia, reglas inconscientes (cf. Johansson, et al., 2005).

Ahora bien, si aceptamos que el conocimiento participa en las dos fases del proceso de la valoración estética, tal y como lo acepta Prinz, ¿son argumentos descalificadores de un papel causal de la razón en la valoración estética decir que el razonamiento nos puede llevar a elecciones que llegamos a lamentar, o que se presenta

de una manera *post-hoc*? El primer argumento es realmente débil, puesto que, más que el razonamiento, las causas frecuentes de elecciones y preferencias que posteriormente lamentamos son las emociones. En cuanto al segundo argumento, por su parte, podemos decir que si bien es cierto que el razonamiento explícito suele ser *post-hoc*, esto no constituye un argumento para afirmar que no pueda darse un razonamiento de manera inconsciente —protológico o no-conceptual— al momento de aplicar reglas en la valoración.

Con todo, Prinz insiste en que la valoración es un proceso afectivo. De acuerdo con él:

Todas las características que hacen buena [good-making] a una obra son agregadas en conjunto y combinadas con características que la hacen mala [bad-making], y el resultado es un nivel total [over-all] de “bueno” [goodness] (o “malo” [badness]), que es lo que reportamos cuando verbalmente apreciamos la obra como buena o mala. Yo propongo que las unidades “buenas”, que son tabuladas de esta manera, son afectivas. Cualquier característica que apreciamos como buena, bien sea consciente o inconscientemente, contribuye con un poco de emoción positiva. Las reglas evaluativas que nosotros aportamos, generan emociones positivas (cf. Prinz, 2007b, 5-6).

Creo que con esto y con lo dicho hasta acá, obtenemos ya una respuesta a las tres preguntas que planteamos para entender qué significa que la apreciación estética tenga un fundamento emocional. Por un lado, hemos visto las dos fases del proceso de la valoración estética que Prinz tiene en mente; fases en las cuales él mismo acepta la participación del conocimiento, pero que, sin mayores fundamentos, considera irrelevante, secundaria e, incluso, nula —en tanto que sería inconsciente—. Por otro lado, también hemos visto que desde el mismo momento de la percepción de la obra, todos los factores quedan supeditados a la asociación, conmoción y afianzamiento de un determinado conjunto de emociones, el cual constituye ya la valoración misma. Esto último en razón de que, de acuerdo con el pasaje anterior, el juicio posterior de

“bueno” o “malo” parece ser tan solo la expresión verbal de dicha valoración, que es ella misma un estado emocional.

Volviendo, entonces, al esquema presentado por Prinz, aquel según el cual si una obra *W* tiene una característica *F*, entonces, en esta medida *W* es buena al grado *N*”, cabe decir que aquí lo “bueno al grado *N*” está constituido por un sentimiento positivo de grado *N*. Según Prinz, “[...] también hay reglas de emoción negativas (correspondientes a características que menospreciamos), que contribuyen a las emociones negativas. Cada característica que evaluamos de esta manera contribuye al estado emocional total que resulta de nuestro encuentro con la obra, y la valencia e intensidad de este estado emocional total normalmente constituyen nuestra apreciación estética” (Prinz, 2007b, 6).

Así pues, hasta aquí la idea de Prinz ha consistido en articular una explicación reduccionista de la apreciación estética, de acuerdo con la cual la apreciación está constituida por una respuesta emocional positiva. El resto del artículo se encamina a la indagación de cuál podría ser, exactamente, esta emoción positiva (cf. Prinz, 2007b, 8). No he considerado pertinente examinar aquí esta indagación; por ello, y sin ir muy lejos, cabe decir que la apreciación estética se reduce —de acuerdo con Prinz— a una emoción que no es propia del dominio de esta apreciación: se trata de una forma de asombro (cf. Prinz, 2007b, 12).

Dicho esto, creo que los argumentos aquí presentados son suficientes para mostrar que, al decir que la apreciación consiste en un estado emocional, Prinz no solamente sugiere que la apreciación sea el resultado de esta respuesta sino que también considera que dicha respuesta es suficiente. Es por ello que su propuesta no contempla una participación igualmente importante del conocimiento en la apreciación estética y



que constituye, así, una tesis unilateralmente reduccionista. Por esta razón, permítaseme pasar ahora a examinar una explicación alternativa, encaminada a mostrar la relación que existe entre el arte y el conocimiento y, a partir de allí, a defender el papel que cumple el conocimiento en la apreciación de una obra de arte. Esto con el propósito, aún más puntual, de mostrar en qué sentido la tesis de Prinz, además de ser unilateralmente reduccionista malentendiende la forma en que funcionan las obras de arte.

## 3.2 ARTE Y CONOCIMIENTO

El problema que subyace a nuestro entendimiento y subsiguiente apreciación de las obras de arte tiene que ver con la referencia. Esto en razón de que, como bien lo señala Catherine Z. Elgin, es “una teoría de la referencia la que debe identificar y caracterizar las relaciones entre un lenguaje (o, más ampliamente, un sistema simbólico) y sus objetos, y explicar las formas en las que el lenguaje funciona en o contribuye a nuestro entendimiento de estos objetos” (Elgin, 1983, 5). Podemos usar el término ‘lenguaje’, sin mayores inconvenientes, haciendo la salvedad de que es preciso distinguir entre lenguajes verbales, propios de la ciencia, y lenguajes no-verbales, tales como diagramas, pinturas y partituras musicales, por ejemplo. Así, puesto que la ciencia, como las artes y el sentido común contribuyen de igual manera a nuestro entendimiento del mundo y de nuestro lugar en él, una adecuada teoría de la referencia debe dar cuenta tanto de lo literal como de lo metafórico, de lo factual como de lo ficticio, de lo descriptivo como de lo expresivo.

### 3.2.1 El lugar del entendimiento en la apreciación del arte: *Un argumento técnico*

Una teoría de este tipo fue la que desarrolló Goodman en su libro *Los lenguajes del arte* (cf. Goodman, 1976). Allí, él expuso ampliamente la denotación y la ejemplificación

como los dos modos básicos de la referencia, para explicar las maneras en que algunos símbolos se refieren entre sí. Aquí, mi argumento de tipo técnico consistirá entonces en señalar, de la manera más clara y sucinta posible, los aspectos principales que de esta teoría son relevantes para nuestro entendimiento y subsiguiente apreciación de las obras de arte; porque, como intentaré mostrar aquí, no podemos apreciar una cosa como obra de arte sin un entendimiento mínimo de la manera en que esta cosa funciona como una obra de arte. Para este propósito, he optado por dejar a un lado el problema de la denotación y me centraré en la ejemplificación, debido a que la expresión —que es propia de las obras de arte— funciona como un caso de ejemplificación metafórica. Aquí, me apoyaré, además, en la sistematización y extensión que del trabajo de Goodman hace Elgin en su texto *With Reference to Reference* (cf. Elgin, 1983).

### 3.2.1.1 La ejemplificación

Comencemos con un breve ejemplo acerca de la manera en que funciona una muestra. Supongamos que usted está interesado en cambiar el color de las paredes de su casa. Para ello, el señor de la tienda de pinturas, a quien usted ha contactado previamente, le lleva a su casa una serie de láminas de diversos colores. Las láminas son pequeñas piezas de acrílico, de forma rectangular y, digamos, fueron hechas en Bogotá. Siendo así, cada lámina instancia el predicado ‘hecho de acrílico’, ‘rectangular’, ‘hecho en Bogotá’ y otros muchos predicados, que incluyen el color de cada una. Pero si tomamos una, por ejemplo, la lámina verde oliva, podemos decir que aunque ella instancia el predicado ‘verde oliva’ y todos los demás predicados, ella solo es una muestra de ‘verde oliva’ y no es, en contraste, una muestra de ‘hecho de acrílico’, ‘rectangular’ o ‘hecho en Bogotá’. Esto en razón de que ser una muestra de ‘verde oliva’ no depende aquí de la instanciación que, de una manera muy visible en comparación con las otras etiquetas o predicados, hace este objeto de la etiqueta verde oliva. Muchas cosas que son

ostensiblemente verde oliva no son muestras de este predicado —tales como los cables eléctricos, por ejemplo—. Y, además, el objeto en cuestión tampoco es una muestra de ‘rectangular’ aun cuando este predicado sea ostensiblemente visible en él.

Esto no debería representar mayores inconvenientes, puesto que si alguien defendiera que la muestra también es una muestra de ‘rectangular’, quizás quedaría un poco confundido cuando, una vez hubiera elegido la muestra verde oliva para pintar su casa, el pintor le cubriera las paredes de la casa con pequeñas piezas rectangulares de color verde oliva. Pero la pregunta que sí es relevante es, entonces, ¿Cuándo algo es una muestra? Esta cuestión la responde Goodman señalando que algo funciona como una muestra cuando funciona como un símbolo de una etiqueta que tiene la capacidad de instanciar (cf. Goodman, 1976, 68). El símbolo se refiere a la etiqueta y actúa así como una representación del conjunto al que se aplica la etiqueta. Para que la lámina verde oliva sea, entonces, una muestra de ‘verde oliva’ debe instanciar la etiqueta ‘verde oliva’ y, además, referirse a ella, esto es, debe funcionar como un símbolo.

Los símbolos, a su vez, están determinados por el uso. Así, por ejemplo, una bebida energética es sencillamente un líquido, pero en este caso ella no está cumpliendo una función simbólica ni referencial. En cambio, en un contexto específico la misma bebida podría ser una muestra de ‘azúcar’, ‘taurina’ o ‘estimulante’; y en muy pocos contextos, ella podría funcionar como una muestra de ‘empacado al vacío’, ‘color café’ o ‘gratis’. Todo depende así de la interpretación que se haga de las muestras; interpretación que está enmarcada, a su vez, dentro de un sistema semántico, puesto que saber si algo funciona como una muestra requiere saber algo acerca del sistema al que pertenece y del papel que cumple dentro del mismo (cf. *ibíd.*).

Así pues, cuando un objeto cumple la doble función de la que he venido hablando, instanciar y referirse a una etiqueta, es cuando decimos entonces que se trata de una muestra o de un ejemplar. Utilizaré esta última palabra, en lugar de la primera, porque no hay una palabra para referirse a la acción y al efecto de mostrar, como sí la hay para el verbo ejemplificar, esto es, 'ejemplificación'. La ejemplificación es, así, una de las maneras en que una cosa puede cumplir una función referencial, y mi interés en este modo de referencia obedece a que la expresión, propia de las obras de arte, es un tipo particular de ejemplificación.

Para mostrar cómo ocurre esto, es decir, el sentido en que las obras de arte funcionan como un ejemplar, permítaseme apelar a un ejemplo de Goodman. Supóngase que estamos frente a un cuadro con árboles y rocas al borde del mar, pintado en grises mates, y con una expresión de gran tristeza. Esta descripción, en palabras de Goodman, ofrece tres tipos de información:

- (1) Las cosas que representa el cuadro;
- (2) Las propiedades que posee; y
- (3) Los sentimientos que expresa (Goodman, 1976, 65).

Así, de acuerdo con los dos primeros casos y en cuanto al carácter lógico de las relaciones en cuestión, se nos dice que: (1) el cuadro hace referencia a una cierta escena; y (2) que es una aplicación concreta de ciertos matices de gris. Sin embargo, el carácter lógico de la relación que el cuadro establece con lo que se dice que expresa, esto es, con el sentimiento de tristeza, ya no parece tan fácil de determinar. Es claro que un cuadro no es triste del mismo modo en que es gris y, por esta razón, Goodman cree necesario establecer un límite entre la 'posesión' y la 'expresión' de una propiedad. Esto obedece a que, en vista de que también es posible describir la situación del ejemplo diciendo que el cuadro es un cuadro triste, o que el cuadro posee tristeza,

se tiene que explicar este nuevo sentido de posesión, de tal manera que nos permita distinguir entre el modo en que el cuadro es triste y el modo en que es gris. En palabras de Goodman, “Un cuadro es gris en tanto pertenece realmente a las clases de cosas que son grises, pero solo metafóricamente posee la tristeza o pertenece a las clases de cosas que se sienten tristes” (Goodman, 1976, 66). Sin embargo, antes de que nos apresuremos a establecer esta distinción en términos de realidad y apariencia o falsedad, en el sentido de decir que lo figurado o lo metafórico implican una posesión no real, Goodman llama nuestra atención sobre el hecho de que si bien lo metafórico verdadero no es literalmente verdadero, tampoco es puramente falso.

Varios aspectos podemos señalar ya con respecto a ciertos problemas en los que parece incurrir la postura emocionalista de Prinz frente a la apreciación del cuadro en cuestión. Por un lado, al defender que la apreciación del cuadro consiste en una respuesta principalmente emocional, el emocionalista bien podría obviar las cosas que el cuadro representa y las propiedades que posee, centrándose así en aquello que el cuadro expresa. Esto en razón de que determinar dichas cosas y propiedades requiere, sencillamente, de un mínimo de conocimiento —cuya aplicación, por lo demás, debe ser racional—. Pero aun suponiendo que el proceso de apreciación emocional sí atiende a estas cosas —como sugiere Prinz—, no es claro en qué sentido el conocimiento aporta a la apreciación pero solo la emoción la configura y, finalmente, la constituye. Así, tal y como vimos al exponer la postura de Prinz, el emocionalismo sugiere que esta participación del conocimiento, que tiene lugar en las dos fases de la apreciación, es una participación inconsciente y que por ello pierde su estatus racional.

Todo el carácter reduccionista y unilateral de la tesis emocionalista parece descansar, entonces, sobre este punto: Prinz acepta que tanto el conocimiento como las emociones participan en el proceso de apreciación estética, pero también afirma que la

participación del conocimiento es inconsciente —no racional— y que, además, es una participación que se encuentra en función de un estado emocional. De manera implícita, Prinz parece suponer, entonces, que las emociones tienen un fundamento protológico, no conceptual y que, en consecuencia, son suficientes para el proceso de la apreciación estética. Este supuesto, por una parte, no ha sido demostrado por el autor y, por otra, tan solo oscurece nuestra comprensión del proceso de la apreciación, que bien podría explicarse señalando una participación conjunta —e igualmente relevante— del conocimiento y de la emoción.

Ahora bien, aun aceptando que esto sea posible y que, por ejemplo, el conocimiento opere aquí de una manera no conceptual, el emocionalismo de Prinz tampoco explica satisfactoriamente el sentido en que si bien lo metafórico verdadero no es literalmente verdadero, tampoco es puramente falso. Prinz acoge el relativismo y el subjetivismo que, según él, surge de la estética popular y que tiene fundamentos emocionales. ¿No es posible hablar, entonces, de criterios de corrección en las artes? ¿No podemos escapar a la creencia popular de que en cuestión de gustos no hay disgustos? En mi opinión, ambas preguntas pueden obtener una respuesta satisfactoria: podemos hablar de criterios de corrección en las artes y podemos escapar al subjetivismo.

Así pues, antes de entrar a aclarar el sentido metafórico de la expresión, el cual explica estas cuestiones que la tesis emocionalista deja en el misterio, permítaseme hacer un comentario adicional sobre la manera en que funciona la ejemplificación, en contraste con otro tipo de referencia, como lo es la denotación. Me refiero a la distinción entre la ‘dirección’ y el ‘dominio’ de la denotación y la ejemplificación. De acuerdo con esta distinción, la representación es un problema de denotación, en cuanto denotar es hacer referencia a algo (el símbolo, predicado o etiqueta se dirige,

desciende, hacia el objeto que representa), pero la ejemplificación es, en cierto modo, un problema de posesión que va en dirección opuesta a la denotación (el objeto mismo, que funciona como un símbolo, se remonta hacia lo denotado por la etiqueta). Así, por ejemplo, decimos que un cuadro es gris, si, y sólo si, ‘gris’ se aplica a él, es decir, si es denotado por la etiqueta ‘gris’; pero decimos que un cuadro ejemplifica el color gris, si, y sólo si, el cuadro mismo se toma como una muestra, al igual que la muestra del pintor, que soporta el rango de objetos a los cuales la etiqueta ‘gris’ se aplica. Aquí, la diferencia de dirección radica en que la ejemplificación es un modo de denotación invertida. En cuanto a la diferencia de dominio, esta se refiere a que mientras cualquier cosa puede ser denotada por una etiqueta, solo las etiquetas pueden ejemplificarse.

Esta distinción nos sirve, además, para entender una relación adicional que hay entre la ejemplificación y la expresión, pues si bien es cierto que un cuadro puede ejemplificar la etiqueta ‘gris’, no por ello podemos afirmar, propiamente hablando, que el cuadro expresa el color gris. Esto sucede, de acuerdo con Goodman, porque “no toda ejemplificación es expresión, pero toda expresión es ejemplificación” (Goodman, 1976, 67). Para entender esto, debemos entrar a examinar la función simbólica de la ejemplificación metafórica, esto es, el sentido metafórico de la expresión.

### 3.2.1.2 Ejemplificación metafórica

Ya vimos que un cuadro es gris en tanto pertenezca realmente a la clase de cosas que son grises, y que solo metafóricamente posee la tristeza o pertenece a la clase de cosas que se sienten tristes. Es de anotar que, en términos generales, tanto el juicio de que un cuadro es gris como de que es triste son ambos informes de lo que el cuadro ejemplifica. Al decir que un cuadro ejemplifica literalmente el gris y metafóricamente la tristeza, nos encontramos frente a un caso de *transferencia* de etiquetas, puesto que la

etiqueta 'triste' ha sido transferida desde el reino de los 'objetos que expresan sentimientos' hacia el reino de los 'objetos coloreados'. Así, de acuerdo con Goodman, la ejemplificación metafórica es la aplicación de un predicado familiar a un objeto nuevo: "Al parecer, la metáfora es algo así como enseñar nuevas artimañas a una palabra vieja, aplicar una vieja etiqueta de una manera nueva" (Goodman, 1976, 83).

En este sentido, como la etiqueta transferida es familiar, ella implica una historia que está en desacuerdo con el nuevo uso, pero, a su vez, se requiere una cierta atracción que permita legitimar este nuevo uso: el uso transferido de esta etiqueta debe poseer la tensión suficiente para señalar la novedad y la consonancia suficiente para adquirir sentido. A este respecto, Goodman señala que "Donde hay metáfora, se da conflicto [...] es un asunto entre un predicado con un pasado y un objeto que cede, aunque con reservas [...] La aplicación de un término es metafórica solo si, hasta cierto punto, es contraindicada" (Goodman, 1978, 83-84). Y es este carácter de atracción y resistencia el que permitirá distinguir el uso metafórico tanto de la falsedad simple como del uso literal. Así, si el uso metafórico implica tanto resistencia como atracción, el predicado será simplemente falso si encuentra resistencia sin atracción, mientras que el uso literal se puede entender si se presenta atracción sin resistencia.

De esta manera, cuando se dice que el cuadro del ejemplo es gris simplemente se le está asignando la etiqueta 'gris' y cuando se dice que es triste, se está re-asignando (*re-assign*) la etiqueta 'triste'; pero cuando se dice que es amarillo, se tiene el caso de una mala-asignación (*misassignment*) de la etiqueta 'amarillo'. Esto último también puede explicarse desde el punto de vista metafórico. Así, por ejemplo, cuando se dice que el cuadro es alegre también se está re-asignando la etiqueta 'alegre', pero de un modo erróneo y, por ende, la aplicación metafórica es falsa. Estos criterios de corrección, cabe insistir en esto, requieren de un mínimo de conocimiento; pero en la



tesis emocionalista de Prinz estos criterios o bien se dejan a una cierta intuición por parte del espectador, o bien se explican a partir de una cierta uniformidad en nuestras emociones y, en consecuencia, se abre la puerta así a un completo relativismo estético en el cual todo juicio sería igualmente válido. La explicación cognitiva de Goodman, que también contempla la participación de las emociones, sí permite explicar muy bien la noción de falsedad involucrada en la aplicación metafórica; para ello solo debemos examinar la manera en que funciona un esquema.

#### 3.2.1.2.1 Esquema

Para Goodman, una etiqueta no funciona aisladamente sino vinculada a una familia (cf. Goodman, 1976, 85). Para él, nuestras categorías, así como los esquemas y sistemas de conceptos, que son igualmente agrupaciones de etiquetas, deben verse como conjuntos de alternativas. Un esquema es pues un grupo de etiquetas que sirve para delimitar un grupo de objetos y el agregado de los rangos de extensión de las etiquetas es un reino (*realm*). Esto se puede explicar mediante la siguiente figura (cf. Fig. 2). Aquí podemos ver que las etiquetas ‘amarillo’, ‘rojo’ y ‘gris’ son los miembros de un esquema que define el reino de las cosas coloreadas. De acuerdo con esto, el rango de extensión de ‘gris’, por ejemplo, comprenderá todos los objetos grises, mientras que el reino en cuestión podrá abarcar todos los objetos coloreados. Una re-asignación o transferencia de etiquetas, que dará lugar a una aplicación metafórica, implicará traer etiquetas de otro esquema y, en consecuencia, implicará también un cambio de reino. Así, por ejemplo, podemos hacer la re-asignación de la etiqueta ‘amarillo’, cambiándola por la etiqueta ‘alegre’. En tal caso —esto es, cuando una etiqueta (‘alegre’, por ejemplo) de un esquema determinado (sentimientos) es aplicada a un objeto de un reino clasificado por otro esquema (el de los colores)—, el resultado es un predicado metafórico.

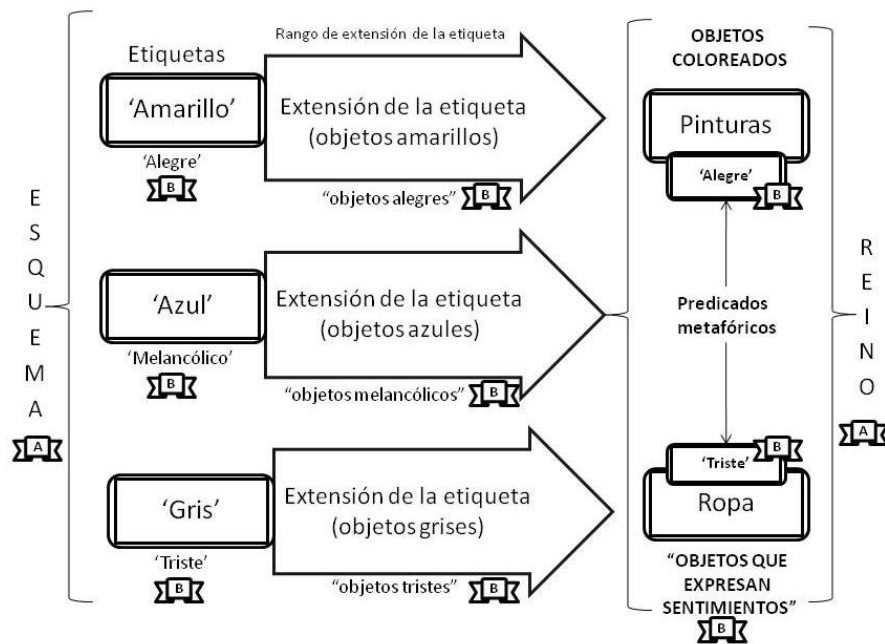


Figura 2.

Nótese también que esta migración de la etiqueta a un reino ajeno siempre está acompañada por la transposición de otras etiquetas del esquema natal. De aquí que las etiquetas no se muevan solas, pues el empleo de las etiquetas del esquema viejo en el reino nuevo debe ser organizado por el empleo tradicional de aquellas etiquetas en su reino de origen. En este sentido, la aplicación del predicado de sentimientos tales como 'alegre', por ejemplo, a un objeto del reino de los objetos coloreados también determina el hecho de que los colores sean organizados bajo otros predicados de sentimientos tales como 'triste', por ejemplo.

De acuerdo con esto, un esquema puede ser metafóricamente transferido a prácticamente cualquier reino, y puede ser aplicado a un reino ajeno de múltiples maneras. El cuadro de un paisaje cualquiera, por ejemplo, puede ejemplificar metafóricamente 'precariedad' como un símbolo estético. Pero el mismo cuadro, si su valor monetario es muy bajo, también puede ejemplificar metafóricamente 'precariedad' como una inversión. Esto significa que, aun cuando toda expresión es una

ejemplificación metafórica, no toda ejemplificación metafórica es una expresión. Para saber si la metáfora, representada aquí por la etiqueta ‘precariedad’, debe ser interpretada de una u otra manera, necesitamos saber desde cuál esquema y sobre qué reino se ha transferido la etiqueta. Así, puesto que la pintura en cuestión involucra tanto el reino estético como el financiero, saber que la metáfora simplemente se refiere a la pintura no es suficiente. En contraste, al identificar la metáfora como una etiqueta estética, ahí sí su aplicación se vuelve determinada.

Sin embargo, como lo señalé anteriormente al introducir la noción de la ejemplificación (cf. *infra*. secc. 3.1.1.1), en este punto nuevamente es necesario contar con una adecuada interpretación del sistema semántico en el cual está inmersa la etiqueta. Es claro que las obras de arte no vienen con un código, previamente establecido, para ayudarnos a comprenderlas como obras de arte —una especie de tabla de instrucciones que diga: “Léase de esta manera...”—. Pero también debería ser claro que nuestra comprensión de ellas está determinada por nuestro conocimiento de su sistema semántico. Así, por ejemplo, si sabemos que, en cuanto símbolo musical, el movimiento final de una sinfonía ejemplifica ‘alegría’, la metáfora nos puede servir como piedra de toque contra la cual probar nuestras conjeturas acerca de cuáles etiquetas pueden ser relevantes, estéticamente, para su comprensión. Enseguida volveré sobre este tema. Por lo pronto, creo que podemos entender muy bien la conclusión a la que llega Goodman con respecto a la expresión. Para él:

Lo expresado es metafóricamente ejemplificado. Aquello que expresa tristeza es metafóricamente triste. Y lo metafóricamente triste es realmente, pero no literalmente, triste; esto es, se da bajo una aplicación transferida de alguna etiqueta co-extensiva con ‘triste’ (Goodman, 1976, 98).

Esto significa, volviendo a nuestro ejemplo inicial, que una metáfora no-verbal articulada mediante el juicio de que un cuadro expresa tristeza, puede ser explicada

por: *i.* la ejemplificación como el inverso de la denotación, *ii.* la posesión como ejemplificación, y *iii.* la ejemplificación metafórica como posesión transferida. De esta manera, la expresión se entiende como la posesión metafórica en el ámbito de la representación no-verbal.

### *3.2.1.3 La expresión*

La expresión es una forma de ejemplificación particularmente importante en las artes, pues con frecuencia decimos que una pintura expresa alegría, un poema expresa sublimidad y que un trío musical expresa nostalgia. Sin embargo, tal y como lo señala Elgin, debemos tener en cuenta que “[...] decir que una obra expresa un sentimiento no es lo mismo que decir que esta evoca este sentimiento en su audiencia” (Elgin, 1983, 82). Muy a menudo, los sentimientos que expresa la obra son muy diferentes a los que el espectador experimenta. Así, por ejemplo, una obra que expresa sufrimiento puede causar compasión en su audiencia, y una que expresa indiferencia puede causar enfado. Más aun, señala Elgin:

[...] una obra de arte no falla en expresar sus sentimientos incluso si, por ignorancia, o falta de atención, o insensibilidad, la audiencia no responde a la obra. Y la obra tampoco necesita expresar el estado mental de su artista. Un actor de teatro puede expresar alegría incluso si, habiendo ejecutado el mismo papel durante años, él esté sufriendo de un profundo aburrimiento. La emoción predominante de Mozart en los últimos trabajos de su vida, si creemos en la historia, fue de desesperación a causa de su lamentable estado financiero. Aun así, sus obras no expresan esta desesperación, y no lo podrían hacer —incluso si se ejecutara ante una audiencia afligida por las mismas razones— (ibíd.).

Esto significa que las etiquetas metafóricas describen la obra como tal y no al artista o a su audiencia —como bien se podría inferir de la tesis emocionalista—, puesto que son las características mismas de la obra las que sugieren la aplicación de estas etiquetas. Sin embargo, Elgin se pregunta ¿bajo qué circunstancias, entonces, un objeto expresa tales etiquetas? Ella piensa que si bien es cierto que las etiquetas

expresadas son metafóricamente ejemplificadas, no cualquier ejemplificación metafórica lo hará, puesto que un objeto expresa solo aquellas etiquetas metafóricas que ejemplifica como un símbolo estético (cf. Goodman, 1976, 100). Así, por ejemplo, un mismo cuadro ejemplifica metafóricamente ‘sublimidad’ en cuanto obra de arte, ‘un fondo para el retiro’ en cuanto inversión y ‘un boom’ en cuanto contribución a la decoración. Esto significa que, al mismo tiempo, el cuadro funciona metafóricamente en muy diversos sistemas simbólicos. Pero el cuadro únicamente expresa ‘sublimidad’, porque en su ejemplificación de las otras dos etiquetas no funciona como un símbolo estético.

La expresión, de acuerdo con esto, es propia de las artes, pues es justamente nuestra atribución de expresividad al cuadro la que contribuye a nuestro entendimiento y apreciación del mismo como una obra de arte, si y sólo si la metáfora funciona en él como un predicado estético. De aquí que la comprensión de una obra de arte requiera de una adecuada interpretación del sistema semántico en el que está inmersa —pues de no contar con esta interpretación, como nula u oscuramente cuenta el emocionalista, es muy probable que ni siquiera la entendamos como una obra de arte—.

Ahora bien, en la cotidianidad concebimos que todos los predicados emotivos se utilizan para expresar, pero la expresión misma, tal y como la hemos expuesto aquí, no se restringe al uso de estos predicados o etiquetas. Un cuadro puede expresar ‘alegría’ o una escultura ‘sublimidad’, pero ya vimos que hay una restricción sobre las metáforas que pueden ser expresadas por un símbolo estético: únicamente aquellas que son importadas de un reino ajeno cuentan como expresadas (cf. Goodman, 1976, 86). Así, por ejemplo, podría ser completamente lícito decir que una actuación ejemplifica ‘movimiento sutil’, pero no por ello podemos afirmar que el cuadro expresa esa etiqueta. Esto en razón de que la metáfora, en este ejemplo, resulta de re-aplicar un

esquema a su propio reino. En contraste, un poema sí podría expresar un ‘movimiento sutil’, puesto que esta etiqueta ha sido importada del reino de los gestos y aplicada metafóricamente al de los textos.

No obstante, la pregunta parece seguir en pie, pues aun sabiendo que un símbolo metafóricamente ejemplifica una etiqueta de un reino ajeno, ¿cómo podemos decir que este símbolo expresa esta etiqueta? Sabemos que esto pasará únicamente si la metáfora funciona estéticamente, pero el problema es ¿cómo podemos decirlo? Para ello, Goodman ha sugerido cinco síntomas de lo estético:

(i) ‘densidad sintáctica’, síntoma según el cual las más sutiles diferencias en ciertos aspectos constituyen una diferencia entre los símbolos (propia de sistemas no lingüísticos, como un termómetro no graduado)... (ii) ‘densidad semántica’, síntoma en el cual los símbolos se constituyen de cosas distinguidas por las más sutiles diferencias en ciertos aspectos (como el español cotidiano, o las muchas diferentes palabras mediante las cuales los esquimales se refieren a la nieve)... (iii) ‘repleción relativa’, síntoma en el que, comparativamente, muchos aspectos de un símbolo son significativos (como en los lineamientos de un retrato donde cada característica de forma, línea, etc., cuenta)... (iv) ‘ejemplificación’, aquel en el que un símbolo, aun cuando denote o no, simboliza funcionando como una muestra (como un sonido musical que expresa, por ejemplo, tristeza)... y (v) ‘referencia múltiple y completa’, en el cual un símbolo realiza muchas funciones referenciales integradas e interactivas, algunas directas y algunas mediadas mediante otros símbolos (como en una actuación donde un símbolo se emplea en varias capacidades referenciales) (cf. Goodman, 1976, 253-256).

Al afirmar que son síntomas, Goodman no quiere decir que está ofreciendo definiciones, criterios, o guías infalibles de lo estético. El español, por ejemplo, es semánticamente denso y los termómetros no graduados son sintácticamente densos, pero ninguno de ellos se toma como estético. Para él, “un síntoma no es una condición ni necesaria ni suficiente, simplemente tiende, junto con otros síntomas semejantes, a darse en la experiencia estética” (Goodman, 1976, 255). Es decir, Goodman no sugiere que los síntomas puedan ser disyuntamente necesarios o conjuntamente suficientes. Puede ser el caso que, para funcionar estéticamente, un símbolo necesite exhibir por lo menos uno de los cinco síntomas y que cualquier símbolo que los exhiba a todos tiene

grandes posibilidades de funcionar estéticamente. Sin embargo, Elgin enfatiza en que no podemos perder de vista que esto no sugiere que un símbolo sea más estético entre más síntomas exhiba, puesto que el uso del término aquí es apóposito (cf. Elgin, 1984, 83). Así, por ejemplo, un paciente puede exhibir pocos síntomas de una enfermedad incluso si está sufriendo de un caso severo, mientras que otro puede mostrar muchos de sus síntomas y tener, sin embargo, un caso tenue. En consecuencia, Goodman no está ofreciendo pruebas decisivas para saber si un símbolo funciona estéticamente y tampoco para saber si una metáfora ejemplificada es expresada.

La idea de esta teoría de los símbolos, aplicada al arte, es más bien enfatizar en que la caracterización de un símbolo como expresivo es relativa a su identificación como un símbolo estético. Esto es de vital importancia para la apreciación del arte, puesto que de aquí se sigue que no podemos apreciar una obra de arte si no la identificamos como una obra de arte y, más aún, si no entendemos sus caracteres propios que la hacen diferente de otras obras de arte. Si esta doble tarea la podemos realizar única y exclusivamente por medio de una respuesta emocional, de acuerdo con la postura de Prinz, entonces tendremos que aceptar que las emociones tienen un componente cognitivo, aun cuando éste no sea racional —sino, por decirlo de alguna manera, protológico o no conceptual—. De acuerdo con esto, aunque la ‘ejemplificación’ haya sido introducida aquí como un término técnico, es de anotar que es esta conclusión una forma de demostrar que este término es de interés estético, en tanto que provee una explicación de un significativo cuerpo del discurso estético, tal y como lo es la expresión. La expresión, por su parte, no es un término técnico y, por el contrario, nace en el seno del vocabulario de la estética. Por esta razón, la explicación de la expresión en términos de la ejemplificación clarifica la estructura semántica de un

cuerpo de la explicación estética y demuestra la insuficiencia de un acercamiento primordialmente emocional a dicho cuerpo.

### 3.2.2 El lugar del entendimiento en la apreciación del arte: *Un argumento intuitivo*

Con respecto a la refutación de una explicación de la apreciación estética fundamentada únicamente en las emociones, expondré ahora un argumento, menos técnico y más intuitivo que el anterior. Espero subsanar la extensión del primer argumento con la brevedad del que expondré aquí. Para ello, permítaseme idear la siguiente situación. Imagine a una mujer llamada Xiomi, que cuenta con altos niveles de respuesta emocional, pero con muy bajos niveles de respuesta racional hacia las obras de arte. Imagine que Xiomi va a una exposición de Miró y que queda encantada con el derroche de colores. ¿Son suficientes sus reacciones positivas, perceptuales y emocionales, ante la obra para que ella sepa que le gusta el arte de Miró en cuanto obra de arte? ¿Puede ella preguntarse si realmente está haciendo un juicio estético sobre la obra de Miró? La respuesta es obvia. Xiomi puede preguntarse esto. Ella podría estar completamente insegura de que su juicio es estético, aun cuando la obra le suscite reacciones emocionales y perceptualmente positivas.

Ahora, imagine usted que María, una profesora de filosofía del arte, conoce a Juan, un apuesto agente de viajes. Después de varias invitaciones, ellos entablan una bonita relación sentimental. Él es muy perspicaz e inteligente, y aunque tan solo terminó la secundaria, encuentra una cierta atracción por el oficio de su novia. Así, después de varias peticiones por parte de él, ella decide llevarlo a una de sus conferencias de filosofía del arte. El tema es un tanto técnico, pues gira en torno al problema de la ejemplificación en el arte. Al principio, él presta bastante atención a lo



que dice su novia, pero aunque entiende todas las palabras no logra seguir el discurso. Él, sin embargo, se encuentra emocionado; el tono y la forma con la que hablan su novia y los asistentes, y la manera en que se encuentran vestidos los allí presentes, le resultan ciertamente fascinantes. Posteriormente, María pregunta a Juan por sus impresiones de la conferencia. Él dice que le gustó mucho y lo dice con toda honestidad. Pero cuando ella le pregunta por el tema expuesto, él no puede dar razón, puesto que, por supuesto, su honestidad apuntaba a su gusto por detalles ajenos a la conferencia misma. Juan queda un tanto triste.

Posteriormente, María invita a Juan a una exposición de arte. Nuevamente él se encuentra entusiasmado, pero no quiere repetir la situación posterior a la conferencia. Por esta razón, Juan realiza un previo y arduo estudio del artista que se expone en el museo. Una vez en la exposición, María comenta a Juan sus reacciones emocionales con respecto a las obras, pero él, ya muy experto en el asunto, le responde siempre con observaciones positivas y negativas acerca de la técnica del autor. María reprocha a Juan por no disfrutar de la exposición y él, un tanto angustiado, le replica exaltando el buen manejo que del punto de fuga hace el artista. María queda decepcionada ante la insensibilidad de su novio y él, más que triste, ahora ha quedado bastante confundido.

¿Cuál es mi punto? Pienso que la mayoría de las personas fallan en apreciar las obras de arte, en cuanto arte, porque fallan en entender la manera en que las obras funcionan como arte. En mi opinión, una persona que inicialmente establece un vínculo ‘emocional’ con una obra de arte y se queda en él, como es el caso de Xiomi, nunca podrá apreciar una obra de arte, en cuanto arte, pues si su entendimiento no cumple ningún papel en dicha apreciación ¿cuál sería la diferencia para ella entre apreciar una exposición de Miró y apreciar una piscina de bombas o mirar un colorido álbum de fotografías? Sospecho que ninguna. Xiomi puede decir que le gustó aquello que vio en

la exposición de Miró, pero no que le gustó el arte de Miró. En el segundo experimento mental, el caso de Juan es un tanto diferente. El problema de Juan no es que se quede en una impresión emocional, como es el caso de Xiomi, sino que no sabe dirigir sus facultades emocional y racional, y mucho menos combinarlas. En su primera salida, la situación es clara: Juan no sabe nada de filosofía del arte y, en consecuencia, nunca podrá decir que aprecia o desprecia una conferencia de arte mientras no comience a saber del tema. La segunda invitación, por su parte, ya no parece tan evidente, aunque también debería serlo. En este caso, Juan no sabe apreciar obras de arte, pues su acercamiento a ellas es de tipo meramente racional y no estético. Esta es la situación, por ejemplo, de muchas personas que comercian con arte, pues aunque pueden reconocer y emitir juicios acertados sobre la maestría y el valor de una obra, lo hacen desde un punto de vista mercantil —porque se mueven en ese ámbito—, pero bien puede ser el caso que fallen, propiamente hablando, en la apreciación de la obra desde el punto de vista estético.

### 3.3 UN ARTE APASIONADO PERO INFORMADO

Eileen John, en su artículo “Art and Knowledge”, señala que el debate central acerca del arte y el conocimiento concierne a si el arte puede o no ser una fuente de conocimiento (cf. John, 2005, 329). Como ella, encuentro que afirmar que, en efecto, el arte es una fuente de conocimiento es una cuestión evidente e incontrovertible entre los no filósofos, pero no ocurre lo mismo dentro de la estética filosófica y mucho menos en la filosofía en general, quizás porque no se encuentra una respuesta contundente al ‘qué’ y al ‘cómo’ del conocimiento que transmiten las obras de arte.

John señala dos puntos extremos que parecen ser igualmente tentadores en la reflexión sobre el arte como una fuente de conocimiento. Por un lado, se acepta el arte

con entusiasmo, pero ligeramente más bien como una fuente de comprensión y conciencia fresco. A veces, este enfoque incluye la opinión de que esta comprensión especial no se puede poner en palabras, pero quizás nos permite percibir el mundo de una nueva manera. Por otro lado y en contraste, el arte o la experiencia con el arte es rechazada como carente de los requisitos para la producción del conocimiento, pues este es definido aquí como verdad y creencia justificada. El arte puede ser criticado así por no hacer valer o transmitir creencias verdaderas, o por no proporcionar una justificación para cualquier creencia que pueda transmitir. A veces, este enfoque incluye el cargo de que incluso si afirmaciones sobre conocimiento verdadero son ocasionalmente presentadas en una obra de arte, estas afirmaciones son poco interesantes en contenido.

A este respecto, considero que debemos volver nuestra atención sobre el argumento de tipo técnico. Podemos notar que, con frecuencia, hay un amplio acuerdo entre los críticos sobre la etiquetas que expresan las obras de arte, incluso aun cuando estas etiquetas claramente no sean literalmente verdaderas de la obra. La interpretación de la expresión como ejemplificación metafórica mediante símbolos estéticos es relevante para esto. Las etiquetas expresadas son metafórica, no literalmente, verdaderas de las obras de arte que las ejemplifican. Pero las etiquetas metafóricas son genuinamente, incluso si no literalmente, verdaderas de las obras instanciadas por ellas. Así los caracteres que una obra expresa son caracteres genuinos de la obra. Y un objeto que ejemplifica una etiqueta, literal o metafóricamente, se refiere a esa etiqueta. Es poco sorprendente entonces que las personas que saben cómo interpretar una obra estén de acuerdo en lo que ella expresa. Pero, puesto que las obras de arte son símbolos que hacen referencia a algo, ellas requieren de interpretación. De aquí que sea poco sorprendente que las personas que no saben cómo interpretar una obra a menudo

estén en desacuerdo sobre lo que ella expresa. Puesto que la expresión es ejemplificación mediante símbolos estéticos, los expertos a quienes nos sometemos son aquellos que saben de arte.

La justificación por restricción de las etiquetas que cuentan como expresadas viene de la estética misma. Estas restricciones reflejan prácticas relevantes y se justifican si las metáforas que las satisfacen son significantes en la interpretación de las obras. El entendimiento y valoración de una obra de arte depende de la propia identificación de las etiquetas estéticas que ella ejemplifica literalmente y expresa. Si fallamos al reconocer las etiquetas que una obra ejemplifica literalmente, y fallamos al reconocer lo que ella expresa, fallamos al entender la obra y, en consecuencia, fallamos en la apreciación de la misma. No puede ser cierto entonces que la apreciación estética, tal y como Prinz lo defiende, consista primordialmente en una respuesta emocional.

En suma, he mostrado aquí que la idea de Prinz, al afirmar que la apreciación estética es un estado emocional, es la de defender que la respuesta emocional es suficiente para que surja la valoración estética. En el examen de esta tesis, también espero haber mostrado que Prinz acepta la participación del conocimiento en el proceso de la apreciación, pero la muestra como una participación inconsciente y supeditada a las emociones. En mi opinión, el carácter reduccionista y unilateral de la tesis emocionalista descansa sobre la negación del estatus propiamente racional que tiene el conocimiento en el proceso de la apreciación. De manera implícita, Prinz parece suponer, entonces, que las emociones tienen un fundamento protológico, no conceptual y que, en consecuencia, son suficientes para el proceso de la apreciación estética. Este supuesto, como señalé aquí, no ha sido demostrado y tan solo oscurece nuestra comprensión del proceso de la apreciación, que bien podría explicarse señalando una participación conjunta —e igualmente relevante— del conocimiento y

de la emoción. No es necesario entregarnos al relativismo y al subjetivismo estéticos que se siguen de la postura de Prinz, siendo que aceptando los criterios de corrección estipulados por Goodman bien podemos articular una teoría emocional y racional de la apreciación estética. Al insistir en su tesis reduccionista, pienso que Prinz malentiende, entonces, el funcionamiento de las obras de arte, en cuanto arte<sup>5</sup>.

Así pues, de manera similar a como la hace Prinz, no he afirmado aquí que las obras de arte expresan emociones o, incluso, que ellas necesariamente suscitan emociones, así como tampoco he afirmado que las obras de arte expresan conocimientos o, incluso, que ellas necesariamente suscitan procesos racionales en el espectador. Mi punto principal ha consistido en defender que las obras de arte funcionan, en esencia, como una forma de ‘expresión’, de ‘expresión metafórica’, en particular; y de aquí que una respuesta emocional no sea suficiente para apreciar estéticamente una obra de arte, pues la apreciación misma requiere, además, del entendimiento de la manera en que funciona la expresión metafórica, esto es y sin que sea necesario ser un experto en teoría de los símbolos, entender el sentido en que la obra de arte funciona como obra de arte.

---

<sup>5</sup> Una buena ilustración de las nefastas consecuencias que, en la práctica, conlleva la defensa de este reduccionismo unilateral, puede encontrarse en la divertida situación que Goodman idea acerca de la manera en que los extraterrestres conciben la educación y el uso de las bibliotecas (ilustración incluida aquí en el Apéndice 2).

---

# CONCLUSIONES

En este trabajo he tratado de analizar y de evaluar los aspectos centrales de la que aquí he denominado la tesis emocionalista sobre los juicios de valor. Siendo que por esta tesis he sugerido entender la conjunción de dos afirmaciones —i. “[...] las emociones no están simplemente relacionadas con los juicios morales sino que ellas también son, en algún sentido, tanto necesarias como suficientes para dichos juicios” (Prinz, 2006, 29), y ii. “[...] cuando apreciamos una obra de arte, la apreciación consiste en una respuesta emocional” (Prinz, 2007b, 1)—, he revisado estas afirmaciones de manera separada. Así, del componente moral de la tesis emocionalista me he ocupado en el segundo capítulo, mientras que he abordado el componente estético en el tercero. En el primer capítulo, por su parte, he expuesto el proyecto general en el que se enmarca la tesis de Prinz y, por esta razón, he intentado sistematizar allí los argumentos a favor de este proyecto, esto es, del enfoque naturalista-cognitivo de los juicios de valor. Permítaseme volver, entonces, sobre los aspectos centrales de estos capítulos, con el objeto de señalar las conclusiones obtenidas en cada uno de ellos y de intentar articular la que sería una conclusión general de mi trabajo.

Tal y como lo defendí en el primer capítulo, encuentro muy acertada la postura de Prinz en cuanto a que uno de los preceptos fundamentales que debe tener una filosofía propiamente cognitiva, si es que ha de tener alguno, es la idea de perseguir las explicaciones acerca de los hechos naturales con medios igualmente naturales. De aquí

se sigue que en los ámbitos de la teoría de los valores no podamos aceptar explicaciones de tipo supra-natural, por la sencilla razón de que estos ámbitos son naturales, se derivan de nosotros y están, en consecuencia, anclados en la naturaleza humana y no en un reino espiritual. Estas explicaciones se deben buscar, por tanto, desde un enfoque naturalista-cognitivo de los juicios de valor, enfoque en el que se sostiene la tesis de Prinz.

Para mostrar mejor las fortalezas de este enfoque —en particular, en lo que tiene que ver con su aplicación moral— me pareció pertinente examinar allí la manera en que, según Prinz, podemos defenderlo frente a sus principales contraargumentos. Estos vienen de las críticas a las propuestas naturalistas contemporáneas de la ética, que suelen apoyarse en la distinción ‘es/debe’ o ‘Ley de Hume’, y en las objeciones de G. E. Moore, conocidas como la ‘falacia naturalista’ y el argumento de ‘la pregunta abierta’. Así, por un lado y sin violar los preceptos naturalistas, vimos cómo Prinz idea un sugestivo argumento que preserva y, al mismo tiempo, impugna la ley de Hume, esto es, la imposibilidad de derivar un ‘debe’ de un ‘es’. Por otro lado, también vimos que el autor desvirtúa las objeciones de Moore, mostrando cómo estas incurren en un error categorial, al pretender inferir tesis ontológicas de tesis semánticas.

Siendo estos los principales argumentos en contra de un enfoque naturalista de la moral —que bien podrían extenderse, *mutatis mutandis*, al ámbito de la estética—, creo que una respuesta a ellos tan coherente y fundamentada, como la que expone Prinz, apoya la viabilidad que otras propuestas de la ética, no naturalistas y no cognitivas, podrían encontrar en este enfoque. Tal y como lo hace Prinz, pienso que no se trata de sugerir que estas propuestas se adhieran completamente al enfoque naturalista y abandonen, así, sus formas de indagación; más bien, la idea es propender por la convergencia y el apoyo mutuo. Yo creo que en esta convergencia los límites rígidos e insoslayables entre el trabajo filosófico y el científico se vienen abajo, y aflora

así la necesidad del trabajo interdisciplinario; porque no es lícito que ningún tipo de información y enfoque metodológico puedan ser deslegitimados de manera *a priori*.

Sin embargo, una vez que Prinz pretende poner en práctica las bondades que conlleva este enfoque, parece que olvida justamente este aspecto: la bondad principal que representa dicha convergencia. Esto en razón de que el autor parece sumergirse por completo en una muy sugestiva —pero unilateral y reduccionista— evidencia de la psicología empírica, que le permite apoyar que los juicios de valor tienen un fundamento, principalmente, emocional. El enfoque naturalista de Prinz se pone al servicio, así, de una obstinada defensa de la que aquí he denominado la tesis emocionalista de los juicios valor.

Así, si bien es cierto que hay evidencia empírica a favor de esta tesis y que esta tesis provee elementos de juicio para una sugestiva explicación de cómo ocurren los juicios de valor, debemos tener en cuenta que dicha evidencia no es concluyente y que dicha explicación no goza de una total completitud. Más aún, debemos tener en cuenta que también hay evidencia a favor de la tesis contraria —esto es, a favor de que dichos juicios sean producto de procesos racionales—, y que la explicación que de esta tesis se deriva cubre los vacíos que, justamente, no logra llenar el emocionalismo.

Es por esta razón que, en los capítulos dos y tres, mi trabajo ha consistido en la evaluación de la tesis de Prinz a la luz de las explicaciones alternativas. Mi propósito, sin embargo, no ha sido el de defender la tesis contraria, esto es, la tesis racionalista, sino más bien el de hacer fuerte la viabilidad y necesidad de una conjunción de las dos tesis; esto con la idea de proporcionar una mejor comprensión de la naturaleza de nuestros juicios de valor.

En el segundo capítulo he mostrado el debate actual entre emocionalistas y racionalistas con respecto a la naturaleza de los juicios morales. Allí, espero haber puesto en evidencia que el uso de métodos empíricos para explorar las preguntas



tradicionales en la teoría moral todavía está en su etapa de gestación y que queda mucho por aprender. Pienso que la conclusión principal que podemos obtener de este capítulo es que, *stricto sensu*, el antagonismo entre las posturas racionalistas y emocionalistas no es del todo insalvable. Esto por dos razones, en particular. En primer lugar, vimos allí cómo las posturas de los autores principales del debate surgen de la elección de diferentes situaciones prototípicas: los defensores del enfoque racional-innato se centran en los dilemas morales sofisticados, es decir, sobre aquellos en los que entran en conflicto posturas generalmente aceptadas y que, en consecuencia, hacen más difícil tomar una decisión; por su parte, los defensores del enfoque intuitivo-emocional parecen centrarse en las reacciones que, casi que de manera inmediata, podemos tener frente a las infracciones morales de otras personas e, incluso, de nosotros mismos. De esta manera, la elección de uno u otro dilema como la situación moral prototípica parece arbitraria, y es claro que es ella quien tiene un impacto significativo en el juicio moral resultante.

En segundo lugar y más importante aún, es central el hecho de que unos y otros defiendan los resultados de acuerdo con su postura. Este aspecto opaca, así, al anterior, puesto que si bien es cierto que la elección de la situación moral prototípica parece arbitraria, no se aportaría mayor cosa utilizando una misma situación siendo que las dos posturas interpretarían los resultados a su acomodo —y remitiéndose a agregar, con un extraño tono de modestia, “hay mucho por investigar”—. Así, por ejemplo, y para retomar una de las respuestas de Prinz a Greene, *prima facie*, parece muy cierto que la disminución de la intensidad emocional del método de hacer daño aumenta el índice de aprobación de dicho daño. Pero, *prima facie*, no parece menos cierto que un cálculo que arroje como resultado una disminución en la intensidad del daño mismo aumentará el índice de aprobación de dicho daño. ¿De qué le sirve al debate que el emocionalista contemple factores tales como la cantidad del daño si los termina interpretando, de una

u otra forma, en función de las emociones? Siendo que no le sirve de nada, Prinz se niega a escuchar los planteamientos racionalistas, haciendo que este debate no resulte ‘productivo’.

Volviendo, entonces, a mi conclusión general de este capítulo, pienso que el mejor argumento en este debate es el que busca integrar los dos modelos (junto con la riqueza de la evidencia empírica que cada uno ha generado); aquel que contempla los *inputs* o contextos en los que cada uno tiene mejor aplicabilidad. De esta manera, encuentro que promete mucho más el modelo de Greene, en el que interactúan la razón y la emoción —de acuerdo con las diferentes circunstancias morales que percibe el sujeto—. En mi opinión, un modelo que complemente los procesos racionales con los emocionales puede proporcionar una comprensión más enriquecedora del proceso del juicio moral, puesto que contempla la complejidad de las situaciones morales encontradas en la cotidianidad y sugiere así que, ante dicha complejidad, nuestras respuestas morales no deben ser menos complejas.

Finalmente, en el tercer capítulo también he intentado hacer fuerte esta idea de la conjunción, pero ya desde el ámbito de la apreciación estética. Allí, he mostrado que la idea de Prinz, al afirmar que la apreciación estética es un estado emocional, es la de defender que la respuesta emocional es suficiente para que surja dicha apreciación. Como espero haber mostrado en este capítulo, esta defensa malentende la manera en que funcionan las obras de arte, como un símbolo que refiere de manera metafórica, y no da cuenta de manera satisfactoria de los criterios de corrección en las artes. Por esto, considero que el modelo de Prinz debería ser complementado con una propuesta que, como la de Goodman, dé buena cuenta de estos factores —al contemplar el papel no menos relevante que cumple el conocimiento en la apreciación de las obras de arte—.

Pero, y lo que es más importante aún, en el examen de la tesis de Prinz, también espero haber mostrado que él acepta la participación del conocimiento en el proceso de la apreciación estética, solo que la muestra como una participación inconsciente y supeditada a las emociones. Su tesis descansa, así, en la negación del estatus propiamente racional que tiene el conocimiento en el proceso de la apreciación, puesto que —de manera implícita— Prinz parece suponer que las emociones tienen un fundamento protológico, no conceptual y que, en consecuencia, son suficientes para el proceso de la apreciación estética. Este supuesto, como señalé allí, es indemostrable y tan solo oscurece nuestra comprensión del proceso de la apreciación, que bien podría explicarse señalando una participación conjunta —e igualmente relevante— del conocimiento y de la emoción.

Por esta razón, pienso que en este ámbito sigue siendo válida mi conclusión final al segundo capítulo, pues no le sirve de mucho al emocionalista contemplar la participación de factores diferentes a la emoción si, en última instancia, los interpreta en función de las emociones. En mi opinión, tanto en el ámbito moral como en el estético, la postura de Prinz es, entonces, reduccionista.

En suma, pienso que el enfoque naturalista debe ser implementado en las investigaciones sobre la naturaleza de los juicios morales, pero que esta implementación no debe ser llevada al extremo de la postura emocionalista de Prinz. Este es, básicamente, mi punto de desacuerdo con el autor: su reduccionismo al momento de implementar el enfoque naturalista. Con el planeamiento inicial del naturalismo metodológico, sin embargo, estoy muy de acuerdo en que no se trata de sugerir que las investigaciones no naturalistas se adhieran completamente al enfoque naturalista y abandonen, así, sus formas de indagación. La idea es propender por la convergencia y el apoyo mutuo, pues en esta convergencia los límites rígidos e

insoslayables entre el trabajo filosófico y el científico se vienen abajo, y aflora así la necesidad del trabajo interdisciplinario. No es lícito, como lo señalé anteriormente, que ningún tipo de información y enfoque metodológico puedan ser deslegitimados de manera *a priori*.

Así, por ejemplo, tal y como sugerí en este trabajo, si descubrimos que como consecuencia de un determinado tipo de daño cerebral se ha perdido o deteriorado una determinada habilidad, relacionada con los juicios de valor, mientras que otra aparentemente similar permanece intacta, nos vemos obligados a buscar una diferencia significativa correspondiente a las dos habilidades. Estos hallazgos vienen del trabajo psicológico empírico y la filosofía debe complementarlos mediante su interpretación e incorporación de los mismos en una teoría general. De esta manera, bien harían los teóricos no naturalistas en ética y estética en incorporar las evidencias de la psicología empírica a sus reflexiones, tal y como lo hacen, valga decirlo, la mayoría de los psicólogos con el trabajo filosófico. Ninguna de las dos áreas pierde su estatus con esta forma de trabajo interdisciplinario y, por el contrario, se fortalecen cada vez más en su objetivo común del progreso cognoscitivo. La ciencia, la reflexión filosófica, la práctica, la percepción, y las diferentes artes y moralidades no van por caminos diferentes, sino por una única senda que conduce al conocimiento y a formarse una idea del mundo circundante (cf. Apéndice 2). Nuestro precepto fundamental debe ser, entonces, el de perseguir el entendimiento de este mundo circundante rechazando la idea de que hay cosas que no podemos entender. Esto es lo que caractericé allí, entonces, como un naturalismo metodológico, de acuerdo con el cual las explicaciones acerca de los hechos naturales se pueden hacer con medios igualmente naturales.

# APÉNDICE 1

**Tabla 3: El cerebro moral** (véase también Fig. 3, los colores han sido codificados a esta figura). La primera columna indica ocho áreas del cerebro (áreas de Brodmann (BA) entre paréntesis) relacionadas con la cognición moral mediante estudios de Neuro-imagen. Las columnas posteriores proporcionan información adicional acerca de sus funciones. (Greene y Haidt, 2002, 520-521)

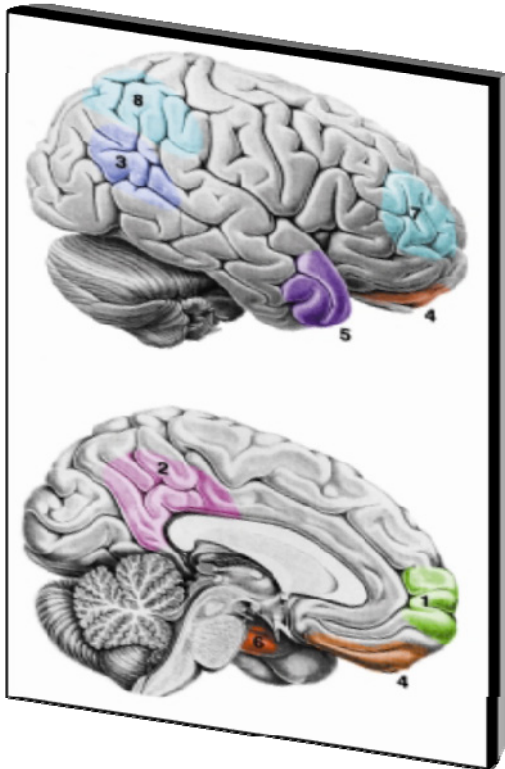
Región del cerebro (con BA)	Tareas morales asociadas	Otras tareas asociadas	Patología social del daño	Funciones probables
1. Giro Frontal Medio (BA 9/10)	Juicios morales personales. Juicios morales impersonales (relacionados con no-morales) [32]. Juicios morales sencillos* [28]. Ver imágenes 'morales' [30]. Juicios de perdón [31]*. (*también frontopolar lateral).	Atribuir intencionalidad al movimiento de formas y caracteres de caricaturas*. Relatos y caricaturas de teoría de la mente [ToM: <i>Theory of mind</i> ]*. Representar estados mentales en un personaje histórico* [36]. Ver rostros de enfado/tristeza [47], imágenes agradables y negativas [48] (con reporte emocional [49]). Retribuir [37]. Ver y/o recordar películas alegres, tristes y repugnantes [50]. Recordar emociones autobiográficas [51]. Planificación emocional [34]. 'Descansar' [42] *(Focalizados en el sulcus paracingulado).	Deficiencia en juicios prácticos [15,16]. Reacciones agresivas [27] y (principalmente en casos progresivos) empatía y conocimiento social disminuidos [18].	Integración de la emoción en la planificación y toma de decisiones [15,16], especialmente en procesos conscientes [33]. ToM [36]
2. Cingulado posterior, Precuneus, Corteza retrosplenial (BA 31/7).	Juicios morales personales. Juicios morales impersonales (relacionados con no-morales). Juicios morales sencillos [32]. [28] juicios de perdón [31]. Imágenes 'morales' [30].	Escuchar episodios autobiográficos afectivos [52] y palabras de amenaza [38]. Leer relatos coherentes, especialmente relatos ToM [53]. Ver caricaturas ToM [54], rostros familiares [55], rostros disgustados, rostros tristes, videos de serpientes, videos de robos previamente experimentados, fotos de combates (e imágenes). Recordar episodios autobiográficos tristes (hombres) [38]. Reconocer palabras neutras de contexto negativo [56]. Planificación emocional [34]. Recordar episodios felices de la vida personal [57], y pares de palabras fáciles de imaginar [39]. 'Descansar' [42].	Deterioro en la memoria de reconocimiento de rostros. Delirio Capgras? (☹) [55].	Integración de la emoción, imágenes (especialmente precuneus [39]) y memoria [38], especialmente para narrativas sociales coherentes.

3. Surco temporal superior, Lóbulo parietal inferior (BA 39).	Juicios morales personales [32]. Juicios morales sencillos [28,29]. Imágenes 'morales' [30].	Ver movimiento biológico (manos, rostros, ojos, cuerpo) [40]; rostros tristes [47]; películas alegres, tristes y repugnantes [50,51]; caricaturas ToM. Leer relatos coherentes con auto-perspectiva y con caracteres, especialmente atribuyendo intencionalidad ToM a formas en movimiento. Representar estados mentales de un personaje histórico [36]. Reconocer palabras neutras de contexto negativo [56]. Recordar pares de palabras fáciles de imaginar [39]. Juicios de interior/exterior Vs. respuesta subjetiva ante imágenes (des)agradables [49]. Ver películas emotivas Vs. Recordarlas [51]. 'Descansar' [42].	Deterioro en el discernimiento de la atención (monos) [40]. Delirio Capgras? (*) [41].	Apoyo en las representaciones de movimientos socialmente significativos [40], y de posibles complejos en las representaciones de la 'personalidad' [41]. ToM [36].
4. Corteza frontal Orbitofrontal/Ventromedial (BA 10/11).	Juicios morales sencillos [28,29]. Imágenes 'morales' [30].	Distinguir recompensa/castigo [37]. Recordar episodios autobiográficos tristes [57]. Reconocer palabras de contexto positivo [56]. Ver rostros enojados [47]. 'Descansar' [42]. (Nota: ausente en muchos estudios PET (**)) de la emoción [34, 48–50]).	Deficiencia en juicios prácticos [15,16]. Reacciones agresivas [27] y (principalmente en casos progresivos) empatía y conocimiento social disminuidos [18]. Dificultad con tareas ToM avanzadas [58].	Representación de valores de retribución/castigo [15,16,37] y control de comportamiento inapropiado/desfavorable [15,27]. ToM 'alto' (***) [58].
5. Polo Temporal (BA 38).	Juicios morales sencillos [28,29].	Leer relatos coherentes (con caracteres) y relatos ToM. Atribuir intencionalidad a formas en movimiento y caracteres de caricaturas. Representar estados mentales de un personaje histórico. Recordar escenas y rostros familiares [36]. Escuchar episodios autobiográficos afectivos [52]. Reconocer imágenes emotivas [59]. Ver imágenes emotivas (con reporte subjetivo) [49], y rostros enojados/tristes [47]. Ver y recordar películas alegres, tristes (ver únicamente) y repugnantes [50]. Recordar emociones autobiográficas [51].	Memoria autobiográfica deteriorada [59].	Impartir un tono afectivo a la experiencia y la memoria [59]. ToM [36].
6. Amígdala	Imágenes 'morales' [30].	Reconocer imágenes emotivas [59]. Ver películas emotivas [51] y rostros tristes [47]. Ver rostros de grupos raciales diferentes [44,45].	Disminución en el juicio social de rostros y movimientos [43].	Rápida valoración de valores de retribución/castigo, especialmente visuales y negativos [43].
7. Corteza dorsolateral prefrontal (BA 9/10/46).	Juicios morales impersonales [32].	Procesamiento de la memoria y otras tareas 'cognitivas' [46].		Procesamiento de la memoria y otras funciones 'cognitivas' [46].
8. Lóbulo parietal (BA 7/40).				

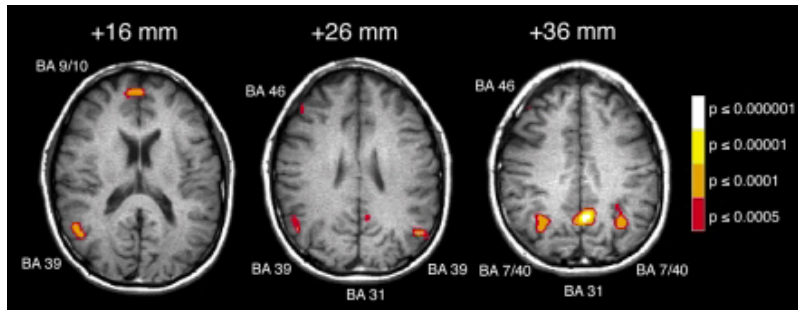
[\*] Engaño de Capgras (*Capgras delusion*): es un fenómeno en el cual una persona cree que un amigo íntimo o un familiar ha sido reemplazado por un impostor de idéntica apariencia.

[\*\*] Estudios PET: Tomografía por emisión de positrones (PET: *Positron Emission Tomography*), es una técnica no invasiva de diagnóstico e investigación por imagen capaz de medir la actividad metabólica de los diferentes tejidos del cuerpo humano, especialmente del sistema nervioso central.

[\*\*\*]ToM ‘alto’: la expresión “‘hot’ ToM” (o “‘hot’ Theory of Mind”) fue acuñada en el área de la neurociencia cognitiva social para referirse a ciertos mecanismos neuronales-cognitivos con los cuales el ser humano responde a señales o intenciones sociales de otros —señales tales como el flirteo o la amenaza—. Los estados mentales o correlatos cognitivos de estos mecanismos son diferentes a los que normalmente se invocan en la teoría de la mente, tales como estados de creencia o conocimiento; para estos últimos se ha acuñado la expresión ‘cold’ ToM (o ‘cold’ Theory of Mind) (cf. Brothers, 2002, 83).



**Figura 3: El cerebro moral** (véase también Tabla 3, los colores han sido codificados a esta figura): Áreas del cerebro relacionadas con la cognición moral mediante estudios de Neuroimagen (áreas de Brodmann (BA) entre paréntesis): 1. Giro frontal medio [*medial frontal gyrus*] (9/10); 2. Cingulado posterior, precuneus y corteza retrosplenial [*posterior cingulate, precuneus, retrosplenial cortex*] (31/7); 3. Surco temporal superior, lóbulo parietal inferior [*superior temporal sulcus, inferior parietal lobe*] (39); 4. Corteza orbitofrontal y corteza frontal ventromedial [*orbitofrontal, ventromedial frontal cortex*] (10/11); 5. Polo temporal [*temporal pole*] (38); 6. Amígdala [*Amygdala*]; 7. Corteza prefrontal dorsolateral [*dorsolateral prefrontal cortex*] (9/10/46); 8. Lóbulo parietal [*parietal lobe*] (7/40). (Greene y Haidt, 2002, 521).



**Figura 4: Áreas del cerebro** (señaladas con las áreas de Brodman (BA)): aquí se exhiben diferencias en la actividad cerebral en respuesta a dilemas morales personales comparados con dilemas impersonales y dilemas no morales [32]. Las áreas que muestran gran actividad para los dilemas personales (comparados con dilemas personales y no morales) son: giro frontal medio [*medial frontal gyrus*] (BA 9/10); giro cingulado posterior [*posterior cingulate gyrus*] (BA 31); surco temporal superior y lóbulo parietal inferior [*superior temporal sulcus, inferior parietal lobe*] (BA 39). Las áreas que muestran gran actividad para los dilemas morales impersonales (comparados con los personales) son: corteza prefrontal dorsolateral [*dorsolateral prefrontal cortex*] (BA 46); y lóbulo parietal [*parietal lobe*] (BA 7/40). Las imágenes son presentadas de izquierda a derecha de acuerdo con la convención radiológica. (Greene y Haidt, 2002, 519).



---

## APÉNDICE 2

Hace unos pocos años, un anciano amigo mío, El Profesor Hans Trublemacher, visitó Marte y escribió un informe sobre el triste estado en que se encontraba la educación científica en aquel lugar donde, al parecer, la educación científica que se ofrecía en muchas instituciones importantes era completamente extracurricular y recibía una consideración, en gran medida, muy parecida a la que recibe la educación artística en muchas universidades importantes de la Tierra: esto es, un medio que proporciona a la gente algo que hacer en su tiempo libre.

(Goodman, Nelson. *De la mente y otra materias*. Madrid: Visor, 1995, p. 266)

### ***Un mensaje desde Marte***

*El profesor Hans Trublemacher, un destacado especialista en educación científica, fue llamado recientemente a Marte, en calidad de asesor, a instancia de los marcianos, quienes se mostraban interesados por el estado de las ciencias en la más importante Universidad del planeta. A continuación, se ofrece una carta en la que se describen sus experiencias.*

«Al llegar, fui recibido calurosamente por el Canciller Eric Cobb, quien me contó que el perfeccionamiento de la educación científica en la Universidad era uno de sus intereses prioritarios. Estaba convencido de que ningún estudiante podría tener una educación madura, sin recibir alguna exposición de las ciencias, de que cierta familiaridad con las ciencias mejoraría la calidad de la mente y de la personalidad, e incrementaría la habilidad de gobernantes, abogados, artistas y demás pilares de la civilización marciana. Además, dado que en Marte se había reducido rápidamente las horas de trabajo y se fomentaba la jubilación anticipada, todos tenían algún conocimiento y habilidad científicos en los que ocupar el tiempo libre. Arremetió contra la tradición ascética que desprecia a las ciencias simplemente porque éstas pueden llegar a proporcionar placer y satisfacción. Pero la financiación presentaba grandes problemas.

A los pocos días de haber iniciado mi visita, descubrí que los cursos oficiales de la Universidad de este lugar se centraban casi completamente en las artes, cubriendo todos sus aspectos. Se ofrecía una amplia variedad de programas de titulaciones en las artes para graduados y no graduados, a los estudiantes que, más tarde, llegarían a convertirse en artistas creadores o activos, o a formar parte de los públicos de las diversas artes. La universidad mantenía compañías de actores profesionales, mimos, bailarines y músicos, y la Facultad estaba compuesta por muchos compositores, pintores, escultores, poetas, dramaturgos, arquitectos y coreógrafos, unos ocupados parcialmente en la enseñanza o en el adiestramiento, otros comprometidos solamente con su propio trabajo. Había suficientes teatros, salas de concierto, estudios de todas las clases, museos y personal técnico especializado.

En contraste, los únicos cursos científicos que aparecían en el catálogo eran los de historia de la ciencia, así como cursos sobre la tecnología en la que se apoyaba uno que otro aspecto de las obras artísticas, como la química de pigmentos, la electrónica aplicada al escenario, la óptica fotográfica o la conservación. No había cursos acerca de ninguna de las ciencias tal y como nosotros las conocemos, ningún programa de estudios para pregrado o posgrado en cualquier ciencia, y ningún trabajo de laboratorio o medios para llevarlo a cabo, excepto en las tecnologías mencionadas.

Después de mi estudio preliminar, tuve una entrevista con el antiguo Decano Christopher Chrysler. Cuando hice notar que la denominación de «Decano de las Ciencias y las Artes» parecía algo anómala, en virtud del papel tan secundario que jugaban las ciencias en la Universidad, dijo que había oído que su homólogo en Harvard tenía el título de «Decano de las Artes y las Ciencias». Se declaró, no obstante, predispuesto favorablemente hacia las ciencias. Cuando pregunté acerca de lo que a mí me parecía un programa completamente rudimentario, en el que las ciencias se comparaban con las artes, señaló que yo había tenido únicamente en cuenta, hasta ese momento, los cursos y actividades oficiales, y que obtendría una opinión diferente de la situación en su conjunto, cuando observase todos los progresos que se estaban llevando a cabo en las ciencias sobre una base extracurricular. El Decano estaba orgulloso de las profundas y variadas actividades científicas que estaban siguiendo, y hacía todo lo posible por alentarlas con el escaso dinero disponible. Además, creía que era mejor que las ciencias no estuviesen sujetas en la Universidad a las limitaciones de los programas oficiales, que la Facultad estuviese firmemente convencida de que la inclusión de las ciencias dentro de los programas con prestigio haría que el nivel disminuyese, que el genio y la destreza científicos no se podrían desarrollar mediante la educación oficial, o ser valorados sobre la misma base que el trabajo realizado dentro del programa artístico y de la Universidad; y, finalmente, creía que los estudiantes estarían tan fascinados por sus proyectos científicos y tan volcados en ellos, que no habría necesidad de ofrecer cursos, calificaciones, obtener reputación o recibir reconocimiento oficial. También dudaba de si podría florecer en la atmósfera de la Universidad un verdadero científico creativo, y estaba convencido de que el dominio de una ciencia no dejaría tiempo suficiente a los estudiantes para llevar a cabo cualquier trabajo normal en la Universidad. Me sugirió que hablase con el profesor Lawrence Vincent, quien había sido, durante muchos años, supervisor de la Facultad del Club de las Ciencias, y asesor de confianza de la administración en materia científica.

A lo largo de los días siguientes, me sentí francamente impresionado por la cantidad y variedad de las actividades extracurriculares sobre las ciencias. En las comunas formadas por el cuerpo de estudiantes había varios clubes: clubes de química, biología, física, astronomía, astrología, numerología, etc. Había un destacado y venerable club dedicado al estudio de los experimentos de William Gilbert sobre el magnetismo y a los experimentos de Robert Boyle sobre los gases; y esta Sociedad Gilbert y Boyle llevaba a cabo cada año algunos de estos experimentos, para el deleite de grandes públicos. En algunas comunas, había tardes libres en las que los estudiantes presentaban su propio material y repetían experimentos famosos. Había un laboratorio grande y bien equipado, principalmente para el uso de los estudiantes, y subvencionado por la Universidad, pero dominado por el más bien pedestre Club de las Ciencias. Sin embargo, las cocinas de los dormitorios de las comunas se convertían, a menudo, después de la cena, en improvisados laboratorios, y algunas de las demostraciones más populares tenían lugar allí.

Sin embargo, en ocasiones determinadas se llevaba a los científicos famosos a la Universidad, en la que pasaban apenas dos o tres días, de tal modo que los estudiantes pudieran verlos trabajar. Y, más recientemente, algunos de los miembros del personal técnico que participaban en el trabajo del Club de las Ciencias, habían estado dando cursos especiales no reconocidos oficialmente; por ejemplo,

un curso de dos semanas sobre química, un curso «comprensivo» de tres semanas sobre física moderna, y seis lecciones sobre metalurgia.

El único caso de curso reconocido sobre ciencia que estaba siendo impartido, era el fruto de un subterfugio ideado por un destacado pianista y Profesor de Música, quien estaba convencido de que cierto trabajo en las matemáticas puras conllevaría un cambio muy grato con respecto al omnipresente énfasis puesto con frecuencia en la interpretación. Presentó subrepticamente un curso sobre álgebra moderna bajo el título de «Aspectos matemáticos de la actuación de grupos musicales».

A continuación, examiné los vastos esfuerzos que estaba haciendo el Gobierno, al parecer, para mejorar el estado de las ciencias en la Universidad. Se había encargado la elaboración de un informe, el cual se recibió después de transcurridos casi dos años. Después de haber tomado muy seriamente en consideración el informe durante un año o algo más de un año, se nombró un Consejo sobre las Ciencias, el cual, después de transcurrir otro año, sugirió una recomendación específica. Sin embargo, éste demostró ser inoperante y fue abandonado, a favor de un sustituto rápidamente inventado.

Se creó una Oficina de las Ciencias y se nombró a un Coordinador para las Ciencias. Pregunté al Coordinador, Mr. Paul Purchase, sobre la función de la oficina y acerca de su propio trabajo. Me explicó sus esfuerzos para introducir cierto orden dentro del caos. La escasez de material científico se estaba convirtiendo en un obstáculo menor, gracias a la planificación y a la distribución propias, de tal modo que no todos pedirían tubos de ensayo al mismo tiempo. El espacio limitado del laboratorio era solicitado, algunas veces, por diversos grupos, y en otros casos permanecía libre. Además, algunos experimentos y demostraciones se programaban, a veces, para el mismo día; y esta circunstancia resultaba especialmente desagradable, cuando cuatro grupos diferentes del campus decidían experimentar, en la misma tarde, produciendo gases nocivos. El Coordinador colaboró en la planificación de los programas; y también estuvo gestionando un nuevo programa con el fin de enviar a los estudiantes a recibir lecciones de química y física, impartidas por profesores de fuera. Y estos esfuerzos, aunque útiles, habían resultado algo periféricos; el Coordinador se había visto seriamente perjudicado por tener que cargar también con el trabajo considerable e incompatible que suponía encargarse, él mismo y personalmente, de algunos de los programas científicos extracurriculares.

Hablé, finalmente, con el Profesor Lawrence Vincent, un miembro de la Facultad, cuya labor había sido muy destacada a la hora de idear y dirigir las actividades científicas de la Universidad. Destacó el hecho de que la universidad se encontrase constantemente bajo presión financiera y que no podía pretender conseguirlo todo. La tarea real de la Universidad era la educación en todos los aspectos relacionados con las artes. Él consideraba que el adiestramiento científico era primariamente vocacional, y pensaba que lo mejor sería dejarlo en manos de las escuelas. Además, dado que la ciencia no era tecnología, era, por tanto, algo parecido a una ocupación inútil, teniendo sólo el valor del entretenimiento; y las actividades científicas en la Universidad, como las actividades atléticas, deberían mantenerse al margen del currículo ordinario. Estaba convencido de que el papel propio de la física se asemejaba mucho al del fútbol. Rechazaba de plano cualquier sugerencia para que la Universidad mantuviese laboratorios y científicos investigadores, comprometidos principalmente con la investigación, por la misma poderosa razón que le hacía mantener teatros, compañías y artistas profesionales.

El amigo que me había invitado a venir a este lugar en calidad de asesor me preguntó, durante un paseo por el puerto espacial, sobre las impresiones que había sacado y mis recomendaciones al respecto. Me vi en la obligación de decirle que, en mi opinión, poco podría hacerse hasta que no se llegase a reconocer más la necesidad básica de llevar a cabo un cambio de actitud, y que las nuevas

ideas eran más necesarias que el hecho de recibir, de nuevo, más dinero, o al menos, anterior a todo esto. Él sugirió cambiar su Universidad por la de Harvard; pero yo estaba deseando regresar a casa, allí donde se reconoce la importancia de las ciencias, y las artes se mantienen en el lugar que les es propio».

(Goodman, Nelson. *De la mente y otra materias*. Madrid: Visor, 1995, p. 254-60)

\*

Más tarde, el Profesor Hans Trublemacher realizó otro viaje, a petición de los propios marcianos, con el fin de examinar sus bibliotecas. Cito a partir de una de sus cartas:

«Hace veinte años, los marcianos no tenían bibliotecas. Enviaron un comité a la Tierra con el fin de examinar nuestras bibliotecas y establecer un sistema semejante a su regreso. Pero al parecer algo anduvo mal por el camino.

»En una biblioteca marciana típica no hay ni mesas, ni pupitres, ni cubículos y, rara vez hay sillas, excepto para los guardias. No hay ni estantes que estén a disposición del público, ni libros que se pongan en circulación. En cada una de las salas de lectura, algunos de los libros más importantes estaban expuestos sobre pedestales independientes, junto a la pared y detrás de una barandilla que impedía a los lectores estar a menos de cuatro pies, pasándose las hojas por control remoto. Frecuentemente, grupos de niños recorrían la sala, mientras un profesor les hablaba acerca de los libros. En las bibliotecas más modernas muchos lectores tenían cajetines electrónicos sujetos con una correa, y descubrí que eran proyectores en miniatura que alquilaban las bibliotecas y que proyectaban una secuencia de diapositivas, justo encima del texto, en el campo visual del lector. No pude determinar si el propósito de éstas era asegurarse de que el lector dispusiera de las imágenes apropiadas que acompañaban al texto, o bien marcar el ritmo de su lectura, o si simplemente le daba algo en que ocupar su mente mientras leía.

«Con todo, encontré muy reveladora la imagen de la gente permaneciendo de pie, leyendo tenazmente un libro a más de un brazo de distancia, mientras las máquinas proyectaban cuadros y los profesores charlaban sobre sus ocupaciones; pero no pude llegar a entender por qué la tienda situada a la entrada estaba haciendo un considerable y rápido negocio vendiendo pequeñas (y, por supuesto, casi ilegibles) reproducciones en escayola de algunos de los volúmenes más populares».

(Goodman, Nelson. *De la mente y otra materias*. Madrid: Visor, 1995, p. 266-67).

---

# BIBLIOGRAFÍA

## BÁSICA

BROTHERS, LESLIE

[2002] *Mistaken Identity: The Mind-Brain Problem Reconsidered*, NY: State University of NY Press.

DWYER, SUSAN.

[1999] "Moral Competence". En: *Philosophy and Linguistics*. Murasugi, Kumiko y Stainton, Robert (ed.). Westview Press.

ELGIN, CATHERINE Z.

[1983] *With Reference to Reference*. Hackett Publishing Company: Indianapolis - Cambridge.

GOODMAN, NELSON.

[1976] *Los lenguajes del arte*. Trad. Jem Cabanes. Barcelona: Seix Barral.

[1995] *De la mente y otras materias*. Trad. Rafael Guardiola. Madrid: Visor.

GREENE, JOSHUA D. ET. AL.

[2001] "An fMRI Investigation of Emotional Engagement in Moral Judgment". En: *Science*, Vol. 293, 2105-2108, Septiembre 14.

[2004] "The Neural Bases of Cognitive Conflict and Control in Moral Judgment". En: *Neuron*, Vol. 44, 389-400, Octubre 14.

[En prensa] "Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie". En: *Journal of Experimental Social Psychology*.

GREENE, J., Y HAIDT J.

[2002] "How (and where) does moral judgment work?". En: *TRENDS in Cognitive Sciences*, Vol.6 No.12, 517-523, Diciembre.

HAIDT, JONATHAN

[2001] "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment". 4, s.l.: American Psychological Association, Inc., *Psychological Review*, Vol. 108, 814-834.

HAIDT, J., BJORKLUND, F., Y MURPHY, S.

[2000] *Moral dumbfounding: When intuition finds no reason*. (Manuscrito no publicado, Universidad de Virginia).

HAUSER, MARC D.

[2006] "The liver and the moral organ". En: *Social Cognitive and Affective Neuroscience Advance Access published*, Octubre 20.

HAUSER, M., YOUNG, L., Y CUSHMAN F.

[2008] "Reviving Rawls' Linguistic Analogy: Operative principles and the causal structure of moral actions". En: *Moral Psychology*. Vol. 2, Cambridge: MIT Press.

HUME, DAVID

[2006] *Moral Philosophy*. Indianapolis: Hackett, 2006.

[1981] *Tratado de la naturaleza humana*. Trad. Félix Duque. Madrid: Orbis.

JOHN, EILEEN

[2005] "Art and Knowledge". En: *The Routledge Companion to Aesthetics*, Eds. Berys Gaut y Dominic McIver Lopes, London: Taylor & Francis, 329-340.

MIKHAIL, J.

[2000]. *Rawls' Linguistic Analogy: A Study of the 'Generative Grammar' Model of Moral Theory Described By John Rawls in 'A Theory of Justice'*. Disertación Doctoral, Departamento de Filosofía, Universidad de Cornell.

MOORE, G.E.

[1997] *Principia Ethica*. Trad.] Adolfo García Díaz y Ana Isabel Stellino. México D. F.: UNAM.

NADO, JENNIFER, KELLY, DANIEL Y STICH, STEPHEN.

[2009]. "Moral Judgment". En: John Symons & Francisco Calvo (ed.) *Routledge Companion to the Philosophy of Psychology*. Routledge, secc. VI, cap. 39.

PRINZ, JESSE J.

[2006] "The Emotional Basis of Moral Judgments". En: *Philosophical Explorations*, Vol. 9, No. 1, Marzo. New York: Routledge, Taylor & Francis, 29-43.

[2007a] *Emotional Constructions of Morals*. New York: Oxford University Press Inc.

[2007b, en prensa] "Emotion and Aesthetic Value". San Francisco: Pacific APA, Marzo 17. Traducido al español por Julián Cubillos, con el amable permiso del autor [En prensa].

RAWLS, JOHN.

[2003]. *A Theory of Justice* (Edición revisada). Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

WHEATLEY T. Y HAIDT, J.

[2005] “Hypnotic Disgust Makes Moral Judgments More Severe”. En: *Psychological Science*, Vol. 16, No. 10, 780-84.

### SECUNDARIA

BECHARA, A., DAMASIO, A., DAMASIO, H., ANDERSON, S. W.

[1994] “Insensitivity to future consequences following damage to human prefrontal cortex”. En: *Cognition*, 50, 7–15.

BECHARA, A., DAMASIO, H., TRANEL, A., DAMASIO, A.

[1997]. “Deciding advantageously before knowing the advantageous strategy”. En: *Science*, 275, 1293–5.

BERTHOZ, S., ARMONY, J. L., BLAIR, R. J. R., Y DOLAN, R. J.

[2002] “An fMRI study of intentional and unintentional (embarrassing) violations of social norms”. En: *Brain*, 125, 1696–708.

BLAIR, R. J. R.

[1995] “A cognitive developmental approach to morality: Investigating the psychopath”. En: *Cognition*, 57, 1–29.

BLAIR, R. J. R., COLLEDGE, E., MURRAY, L, Y MITCHELL, D.G.

[2001] “A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies”. En: *Journal of Abnormal Child Psychology*, 29, 491–98.

CLECKLEY, H. M.

[1941] *The mask of sanity: An attempt to reinterpret the so-called psychopathic personality*. St Louis, Mo.: The C. V. Mosby Company.

DAMASIO, A.

[1994] *Descartes' Error*. Boston, MA: Norton.

[2000] *The Feeling of What Happens*. New York: Basic Books.

[2003] *Looking For Spinoza*. New York: Harcourt Brace.

HARMAN, G.

[2000] "Moral Philosophy and Linguistics". En: Harman, G. *Explaining Value*, 217–25. Oxford: Oxford University Press.

KAPLAN, J. T., FREEDMAN, J, Y IACOBONI, L.

[En prensa] "Us vs. them: Political attitudes and party affiliation influence neural response to faces of presidential candidates".

KENNICK, W.

[1958] "Does Traditional Aesthetic Rest on a Mistake?". En: *Mind*, 67, 317-34.

KOENIGS, M., YOUNG, L., ADOLPHS, R., ET AL.

[En prensa] "Damage to the prefrontal cortex increases utilitarian moral judgments". En: *Nature*.

MIKHAIL, J.

[2002]. *Aspects of the Theory of Moral Cognition: Investigating Intuitive Knowledge of the Prohibition of Intentional Battery and the Principle of Double Effect*. Georgetown University Law Center Public Law & Legal Theory Working Paper No. 762385.

MOLL, J., DE OLIVEIRRA-SOUZA R., Y ESLINGER P. J.

[2003] "Morals and the human brain: A working model". En: *Neuroreport*, vol. 14, No. 3, 299-305.

PIZARRO, DAVID A., BEER, JENNIFER S., Y MONIN BENOIT

[2007] "Deciding Versus Reacting: Conceptions of Moral Judgment and the Reason-Affect Debate". En: *Review of General Psychology*, American Psychological Association, Vol. 11, No. 2, 99–111.

PRINZ, JESSE J.

[2002] *Furnishing the Mind-Concepts and Their Perceptual Basis*. Massachusetts: Massachusetts Institute of Technology.

[2004a] *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.

[2004b] "Which Emotions Are Basic?". En: D. Evans and P. Cruse (Eds.), *Emotion, Evolution, and Rationality*, Oxford University Press.

[2005] "Are Emotions Feelings?" En: *Journal of Consciousness Studies*, 12, No. 8–10, 9–25.

SCHNALL, S., HAIDT, J, Y CLORE G.

[En prensa] "Irrelevant disgust makes moral judgment more severe, for those who listen to their bodies".



SANFEY, A. G., RILLING, J. A., ARONSON, J. K., NYSTROM, L., AND COHEN, J. D.

[2003] “The Neural Basis of Economic Decision Making in the Ultimatum Game”. En: *Science*, 300, 1755–7.

TRANEL, D., BECHARA, A., DAMASIO, A.

[2000] “Decision making and the somatic marker hypothesis”. En: Gazzaniga M.(Ed.) *The New Cognitive Neurosciences*. Cambridge: MIT Press, 1047–61.

WEITZ, M.

[1956] “The role of theory in aesthetics”, *Journal of Aesthetics and Art Criticism*, 15, 27-35.