Revista de Humanidades de Valparaíso, 2022, No 19, 313-348 *Humanities Journal of Valparaíso*
DOI: https://doi.org/10.22370/rhv2022iss19pp313-348 eISSN 0719-4242
Sección Monográfica / Monographic Section

La mente moral: cómo cinco grupos de intuiciones innatas guían el desarrollo de varias virtudes, e incluso algunos módulos, específicos a la cultura[1]

Jonathan Haidt; Craig Joseph

[Traducción al español del texto: Haidt, J. y Joseph, C. (2007). The moral mind: How 5 sets of innate moral intuitions guide the development of many culture-specific virtues, and perhaps even modules. En P. Carruthers, S. Laurence y S. Stich (Eds.) *The Innate Mind*, vol. 3. Nueva York: Oxford, págs. 367-391]

1. Introducción

La moral es uno de los pocos temas académicos que parece tener su propio hechizo protector. Un biólogo no está cegado por su propia naturaleza biológica al funcionamiento mismo de la biología. Un economista no se confunde con su propia actividad económica cuando trata de comprender el funcionamiento de los mercados[2]. Pero los estudiosos de la moral suelen estar sesgados por sus propios compromisos morales. La moral es tan debatida y tan importante que a menudo es difícil dejar de lado la propia humanidad e investigarla de una manera clínicamente desapegada. Uno de los problemas es que el estudio psicológico de la moral, como la psicología misma (Redding 2001), ha estado dominado por investigadores liberales en términos políticos (que nos incluye). Esta falta de diversidad moral y política entre los investigadores ha llevado a una reducción inapropiada del dominio de lo moral a cuestiones de daño / cuidado y equidad / reciprocidad / justicia (Haidt y Graham 2007). La moral en la mayoría de las culturas (y para los conservadores sociales en las culturas occidentales), de hecho, es mucho más amplia, e incluye cuestiones de intragrupo / lealtad, autoridad / respeto y pureza / santidad (Haidt y Graham 2007).

[1] Traducción: E. Joaquín Suárez-Ruíz. Revisión: Pablo López-Silva y Rodrigo Braicovich. Esta traducción cuenta con la autorización tanto de Jonathan Haidt, como así también de Peter Carruthers (editor del volumen en el que se publicó originalmente el texto) y de la editorial original (Oxford University Press).

[2] Por supuesto, biólogos y economistas pueden ser ciegamente leales a ciertas teorías académicas, pero sugerimos que estos sesgos son a menudo manifestaciones de compromisos morales, por ejemplo, las polémicas de Steven Jay Gould.

eISSN 0719-4242 – © 2022 Instituto de Filosofía, Universidad de Valparaíso
This article is distributed under the terms of the
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 Internacional License
CC BY-NC-ND

Este artículo trata acerca de cómo la moral podría ser parcialmente innata, con lo que simplemente queremos decir que está organizada, hasta cierto punto, antes de la experiencia (Marcus 2004). Comenzamos defendiendo una concepción más amplia de la moral y sugiriendo que la mayor parte de la discusión sobre lo innato hasta la fecha no ha sido sobre la moral *per se*; sino que se ha focalizado, más bien, en si la psicología del *daño* y de la *justicia* podría ser innata. Una vez que hayamos presentado nuestra posición de que la moralidad involucra cinco dominios, no dos, dirigiremos nuestra atención a las formas en que esta diversa colección de motivos y conceptos podría ser innata. Consideraremos cinco hipótesis sobre los orígenes del conocimiento y el valor moral, y respaldaremos una de ellas (una forma flexible y generativa de modularidad) como la mejor candidata. Luego, desarrollaremos esta versión de la moral modular, describiendo cómo el "primer borrador" de la moral innata se modifica durante el desarrollo.

Específicamente, vinculamos nuestra visión de lo innato moral con la *teoría de la virtud*, un enfoque clásico que es consistente con las contribuciones de muchas perspectivas modernas. Al hacerlo, estaremos ampliando nuestra exploración de las posibilidades que posee la teoría de la virtud, la cual comenzamos en un artículo anterior (Haidt y Joseph 2004). Ahora bien, no estamos proponiendo que la ética de la virtud sea la mejor teoría moral normativa. Hablamos solo de manera descriptiva, y notamos la existencia de una creciente consiliencia entre los escritos filosóficos sobre la virtud y las emociones, la investigación empírica sobre el funcionamiento moral y la ciencia cognitiva. Dicha consiliencia sugiere que la teoría de la virtud puede producir una comprensión profunda de la arquitectura de la cognición social y moral humana.

En la sección final, discutiremos la importancia de lo narrativo en el funcionamiento moral. En algunos aspectos, esto es otra corrección a lo que vemos como un énfasis excesivo en las concepciones deductivas y calculadoras del valor y la racionalidad, tanto entre los filósofos como entre los psicólogos. Intentamos mostrar, en esta última sección, que un enfoque narrativo de la moral encaja bien con la visión nativista de los "cinco fundamentos" que desarrollamos en la primera parte del artículo, y también ayuda a explicar cómo los fundamentos intuitivos y evolucionados de la moral, son elaborados por la actividad cultural en el complejo y diverso funcionamiento moral que poseen los seres humanos.

2. La moral es muchas cosas

Poco después de que los seres humanos comenzaran a escribir, escribieron sobre la moral. Muchos de los primeros textos morales son grandes listas de leyes y prohibiciones (v.g., el Código de Hammurabi; las partes más antiguas del Antiguo Testamento). Pero a medida que avanzaba la Era Axial (800 a. C. - 200 a. C.), muchas culturas orientales y occidentales comenzaron a desarrollar una psicología de las virtudes más sofisticada. Encontramos discusiones explícitas sobre virtudes, a menudo en el contexto de historias sobre modelos de roles a seguir (v.g., Homero y Esopo en Grecia; el Mahabharata en India). Una característica

CC BY-NC-ND

importante de este enfoque implica que la educación moral se logra moldeando las emociones y las intuiciones, en lugar de dictar fundamentos o principios explícitos. Por ejemplo, hemos heredado la sabiduría de Confucio y de Buda como listas de aforismos y metáforas que producen destellos de comprensión intuitiva.

Una segunda característica de estos enfoques basados en las virtudes es que enfatizan la práctica y el hábito más que el conocimiento y el razonamiento proposicionales. Buda instó a sus discípulos a seguir el Noble Camino Óctuple (un conjunto de prácticas diarias), para alcanzar la perfección moral y psicológica. Aristóteles y Confucio compararon el desarrollo de la virtud con una práctica lenta, necesaria para desarrollar lo que ahora llamamos "virtuosismo" en un instrumento musical (Aristóteles 1941; Hansen, 1991).

Para los antiguos había muchas virtudes, las cuales abarcaban la mayoría de los aspectos de la actividad humana. Las virtudes consistían en excelencias que se esperaba que las personas cultivasen en sí mismas, dependiendo de sus roles sociales y posiciones en la vida. Dos de los más grandes pensadores de la filosofía griega antigua, Platón y Aristóteles, aunque tenían nociones muy diferentes de qué eran las virtudes, llevaron a cabo gran parte de sus investigaciones sobre la ética examinando el concepto de virtud y las virtudes individuales, cuáles eran sus fundamentos y cómo se adquieren.

2.1 La ética del dilema y el gran estrechamiento

La idea de que la moral es un conjunto de virtudes que deben cultivarse mediante la práctica siguió siendo el enfoque dominante en todo el mundo hasta, por lo menos, la Edad Media. Santo Tomás de Aquino siguió a Aristóteles tanto en la ética como en otras cosas, e incluso pensadores islámicos, como Miskawayh y al-Ghazali, tomaron prestadas ideas de Aristóteles para construir sus teorías de la moralidad. Incluso hasta mediados del siglo XX, filósofos y psicólogos influyentes (Dewey 1922; Hartshorne y May 1928) continuaron asumiendo la validez esencial de la teoría de la virtud y, de hecho, basaron sus programas de investigación empírica en el supuesto de que las virtudes eran psicológicamente reales y servían para organizar gran parte de la vida moral.

Sin embargo, las ideas de los filósofos occidentales sobre la moral comenzaron a cambiar en el siglo XVIII. En su mayor parte, las morales basadas en la virtud y la religión se caracterizan por creencias específicas y compromisos sustantivos, ideas 'densas' sobre la naturaleza humana y la sociedad. Con la Ilustración, estos supuestos fueron objeto de un escrutinio cada vez mayor, y los filósofos comenzaron a buscar fundamentos para el juicio moral que no dependieran de creencias metafísicas específicas o identidades de grupo. Lo que MacIntyre (1981) ha llamado "el proyecto de la Ilustración" es el intento de fundamentar la moral en verdades sumamente abstractas, incluso verdades lógicas, y desvincularla (especialmente) de las creencias religiosas. Allí surgieron dos tipos de alternativas que siguen teniendo relevancia en la actualidad: las teorías *formalistas* y las teorías *consecuencialistas*. Las teorías formalistas de

CC BY-NC-ND

la ética, de las cuales Kant es el ejemplo más conocido, definen los juicios morales haciendo referencia a su forma lógica, por ejemplo, como máximas o juicios prescriptivos, más que por su contenido. El status moral de una acción se establece por el tipo de norma que la sustenta. Las teorías "formalistas", en el sentido en que usamos el término aquí, también incluirían la mayor parte de las teorías contractualistas, como las de John Rawls (1971) y Thomas Scanlon (1998), así como también las de Locke, Hobbes y Rousseau. De manera semejante a las teorías estrictamente formalistas, el contractualismo intenta fundamentar (o explicar) los juicios morales postulando la existencia de algo así como relaciones contractuales entre agentes. Aunque este tipo de enfoque parece estar más atentos a las realidades de la naturaleza humana y de los arreglos sociales y políticos, continúa intentando basar la moral en relaciones (en este caso entre individuos) formales (en este caso contractuales).

En contraste, las teorías consecuencialistas, especialmente el utilitarismo, intentan explicar y fundamentar los juicios morales en evaluaciones premorales de las consecuencias de las acciones. Lo moralmente correcto se define, fundamentalmente, como aquello que tendrá las mejores consecuencias (sin embargo, esta frase tan importante es generalmente malinterpretada).

A pesar de sus diferencias (que son grandes), tanto los enfoques formalistas como consecuencialistas de la moral buscan separar el juicio moral tanto como sea posible del desordenado mundo de las prácticas sociales y los comportamientos específicos. El formalismo reemplaza el juicio moral sustantivo con una racionalidad lógica, mientras que el consecuencialismo lo reemplaza con una racionalidad calculadora. Ambos enfoques privilegian la parsimonia: las decisiones morales deben tomarse con respecto a un principio fundacional, como el imperativo categórico o la maximización de la utilidad. A su vez, ambos insisten en que las decisiones morales deben regirse por la razón y la lógica, no por la emoción y la intuición. Y ambos desvalorizan lo particular en favor de lo abstracto.

Los puntos en común entre estas dos perspectivas han llevado a un consenso moderno sobre el alcance de la investigación ética: la moral consiste en resolver dilemas que involucran los conflictos de interés de las personas. El filósofo Edmund Pincoffs (1986) llama a este enfoque moderno "ética del dilema" y lamenta la pérdida del antiguo interés filosófico por la virtud. Mientras que los griegos se centraron en el *carácter* y se preguntaban sobre en qué tipo de persona deberíamos *convertirnos*, la ética moderna se centra en las *acciones*, tratando de determinar cuáles son las que deberíamos *hacer*.

A pesar de esto, la ética del dilema ha continuado floreciendo en la filosofía y en la psicología, donde ha guiado la operacionalización de la moralidad. El método pionero de Lawrence Kohlberg (1969) consistió en el estudio longitudinal de cómo los niños resuelven dilemas morales: ¿debería Heinz robar un medicamento para salvar a su esposa moribunda? La conclusión de Kohlberg fue que los niños mejoran progresivamente en la ética del dilema hasta que alcanzan la etapa más alta, la etapa 5, en la que todas las decisiones se toman con referencia a un principio de justicia universalmente aplicable, autoconstruido y no consecuencialista.

CC BY-NC-ND

Carol Gilligan (1982) desafió las conclusiones de Kohlberg utilizando un dilema diferente. Entrevistó a mujeres que se enfrentaban al dilema de un embarazo no deseado y ofreció un principio superior diferente: el cuidado. Los psicólogos sociales también han operacionalizado la moral como un dilema, poniendo a los sujetos de investigación en situaciones difíciles en las que deben tomar decisiones que ayudarán o dañarán a un extraño (v.g., el estudio del "buen samaritano": Darley y Batson 1973; investigación de la empatía-altruismo: Batson *et al*. 1983; estudios sobre la obediencia: Milgram 1963). Baron (1993) ha declarado que el consecuencialismo es la comprensión normativamente correcta de la moral, y gran parte de la investigación realizada en relación con su enfoque implica presentar decisiones alternativas a los sujetos, donde cada una de ellas posee diferentes costos y beneficios. De hecho, cuando los filósofos morales realizan experimentos, como lo están comenzando a hacer, éstos experimentan principalmente con dilemas como el problema del tranvía y del bote salvavidas, que contraponen preocupaciones utilitarias y deontológicas (Greene *et al*. 2001; Petrinovich, O'Neill y Jorgensen 1993).

Incluso cuando los métodos de investigación no utilizan dilemas *per se*, adoptan la condición límite implícita de la ética de dilema: las cuestiones morales son aquellas que pertenecen a los derechos y el bienestar de los *individuos*. Aquello que concierne a la moral es el ayudar y el herir a las personas. Elliot Turiel, un ex alumno de Kohlberg y una figura importante en psicología moral, codificó esta visión de la moralidad centrada en el individuo en su influyente definición del dominio moral en cuanto:

> (…) juicios prescriptivos de justicia, derechos y bienestar concernientes a la forma en que las personas deben relacionarse entre sí. Las prescripciones morales no son relativas al contexto social, ni están definidas por él. En consecuencia, los juicios morales de los niños no se derivan directamente de los sistemas institucionales sociales, sino de características inherentes a las relaciones sociales –incluyendo allí experiencias que involucran daños a personas, violaciones de derechos y conflictos de reclamos en competencia. (Turiel, 1983, 3)

La delimitación de Turiel del dominio moral parece obviamente válida para muchas personas en las culturas occidentales modernas. Sin embargo, en relación con personas de culturas más tradicionalistas, la definición no capta todo lo que para ellas cae dentro del dominio de lo moral. En otras palabras, la definición de Turiel (afirmamos) es inadecuada como generalización inductiva. Es una definición estipulada que no se corresponde con los hechos empíricos. Cuando el dominio moral se *define* como "justicia, derechos y bienestar", la psicología que emerge no puede ser una verdadera psicología de la moral: solo puede ser una psicología de juicios sobre justicia, derechos y bienestar. Cuando el dominio de la moral se reduce de esta manera, florecen teorías demasiado parsimoniosas de la psicología moral. Por ejemplo: la moral puede explicarse evolutivamente como la extensión del altruismo de parentesco más el altruismo recíproco a grupos más grandes que aquellos en los que evolucionamos. A su vez,

CC BY-NC-ND

desde el punto de vista del desarrollo, la moral puede explicarse como la extensión progresiva de la comprensión del niño de que dañar a los demás (lo que incluye tratarlos injustamente, sin reciprocidad) es malo.

Pero, ¿y si la moral consiste en algo más que daños, derechos y justicia?, ¿qué pasa si estas preocupaciones son parte de una capacidad humana más grande y complicada que no puede ser explicada de forman tan parsimoniosa?, ¿acaso podrían haberse desarrollado prematuramente teorías sobre los orígenes y el desarrollo de la moral?

2.3. El renacimiento de la amplitud

Una de las distinciones más importantes en el estudio de la moral, pero también más problemática, es la que existe entre juicios 'morales' y 'convencionales'. Turiel (y los teóricos del desarrollo cognitivo en general) distinguen estos dos dominios del juicio social sobre la base de la presencia de cuestiones de "justicia, derechos y bienestar". Las reglas morales son aquellas relacionadas con la justicia, los derechos y el daño / bienestar (v.g., no golpear, engañar o robar), y no pueden cambiarse por consenso porque hacerlo crearía nuevas clases de víctimas. Por el contrario, todas esas otras reglas que encuentran los niños (v.g., no llame a los adultos por su nombre, coloque su mano sobre su corazón mientras dice el juramento de lealtad) son cuestiones de tradición, eficiencia o coordinación social que también podrían ser diferentes si los que están en el poder, o la gente en general, optara por cambiarlos.

En las sociedades occidentales en las que las personas aceptan cierta versión del contractualismo como base de la sociedad, esta distinción tiene sentido. Sin embargo, en la mayoría de las culturas el orden social es un orden moral, y las reglas sobre la vestimenta, los roles de género, la comida y las formas de dirigirse al otro son cuestiones profundamente morales (Abu-Lughod 1986; Meigs 1984; Parish 1994; Shweder, Mahapatra y Miller 1987; Hampshire 1982). En muchas culturas, el orden social es también un orden sagrado. Incluso una mirada superficial a los textos religiosos fundamentales revela que, si bien los dioses parecen preocuparse por si nos ayudamos o nos lastimamos unos a otros, les importan muchas otras cosas más. Sería un gran malentendido del judaísmo antiguo, por ejemplo, describir los Diez Mandamientos como una mezcla de reglas morales (sobre no robar, matar o mentir) y convenciones sociales (sobre el Sabbath y las formas prescritas de hablar y adorar). Kelly y Stich (en este volumen[3]), de hecho, argumentan que la teoría del dominio moral propuesta por Turiel y otros es simplemente falsa. Los autores cuestionan la comprensión de las categorías de 'moral' y 'convencional' como dominios psicológicamente distintos, y remiten a su propia investigación para señalar que incluso para algunos asuntos vinculados con el daño, los derechos y la justicia (por ejemplo, azotar a un marinero desobediente), los adultos occidentales juzgan las transgresiones como algo históricamente contingente y dependiente de la autoridad (Kelly *et al*. 2007).

[3] N. del T.: Haidt se refiere al capítulo Kelly, D. y Stich, S. (2007).

CC BY-NC-ND

Como psicólogos culturales, compartimos las preocupaciones de Kelly y Stich. Consideramos el funcionamiento moral como un ejemplo del planteo general de que la cultura y la mente "se inventan una a la otra", para usar la frase de Shweder (1990). En contraste con Kohlberg (por ejemplo), consideramos que es importante comenzar la explicación del funcionamiento moral observando los hechos individuales y culturales sobre el funcionamiento moral, y no con una definición estipulada del dominio de lo moral heredada de los filósofos morales. Este enfoque es más "de abajo hacia arriba", no solo empíricamente sino también conceptualmente. Nosotros asumimos (al menos al comienzo del análisis) que lo que la gente piensa que son sus conceptos morales, son, de hecho, conceptos morales (en lugar de dividirlos en conceptos 'morales' y 'convencionales' desde el principio).

Todas las sociedades humanas generan e implementan normas (Brown 1991). Sripada y Stich (2006, 281) han proporcionado una definición útil de 'norma', a saber, "una regla o principio que especifica acciones que son requeridas, permisibles o prohibidas independientemente de cualquier institución legal o social". Resumiendo su discusión, los autores caracterizan las normas como (a) reglas o principios, (b) con normatividad independiente y que generan (c) motivación intrínseca y (d) actitudes punitivas hacia los infractores. En pocas palabras, las personas esperan que los demás actúen de cierta manera y no de otras, y *se preocupan* si los demás están siguiendo o no estas normas. Consideramos que el primer paso para trazar un mapa del dominio moral de cualquier cultura debería ser, por tanto, enumerar y contar las normas que reciben la mayor atención, ¿sobre qué normas y violaciones de normas habla la gente?, ¿qué normas se rompen y castigan en los mitos y cuentos populares? Cuando las personas rechazan o critican a otros miembros de su comunidad, o cuando expresan su conmoción por las prácticas de otra comunidad, ¿qué normas están involucradas? (ver Cosmides y Tooby, en este volumen[4], para un enfoque similar sobre la definición del dominio moral).

Esta etnografía cuantitativa es difícil, pero varios proyectos de investigación han intentado dibujar mapas de forma experimental. Usando su conocimiento de las normas locales en Orissa, India, Shweder, Mahapatra y Miller (1987) crearon una lista de 39 acciones, algunas de las cuales causaron daño de modo directo, otras involucraron cuestiones de comida, ropa, formas de dirigirse a los demás y otras tradiciones que contarían como convenciones sociales según la definición de Turiel. Los sujetos Oriya de Shweder —adultos y niños— dieron respuestas que revelaron un dominio moral muy diferente de las muestras obtenidas en adultos y niños en Chicago. Los estadounidenses encuestados vieron daños y violaciones de derechos en muchas de las acciones (por ejemplo, un esposo golpea a su esposa por desobediencia) y, en consecuencia, las moralizaron. Los encuestados Oriya, por el contrario, revelaron un mundo moral más amplio en el que las cuestiones de respeto y jerarquía (v.g., la obediencia de la esposa a su marido) y pureza/santidad espiritual (v.g., no comer alimentos espiritualmente contaminantes en momentos prohibidos) parecían ser por lo menos tan importantes como las cuestiones vinculadas con daños, derechos y justicia. Haidt, Koller y Dias (1993)

[4] N. del T.: Haidt se refiere a un capítulo que finalmente no fue publicado en el volumen en cuestión.

CC BY-NC-ND

demostraron más tarde que la diferencia cultural no se debía a "daños ocultos", como Turiel *et al*. (1987) había afirmado. Usando un nuevo conjunto de violaciones de normas inofensivas (incluyendo el uso de una bandera para limpiar un inodoro y tener relaciones sexuales con un cadáver de pollo), Haidt *et al*. descubrieron que tan solo una población universitaria estadounidense de élite limitaba el dominio moral a cuestiones de daño, derechos y justicia. Para otros grupos, particularmente para los grupos de bajo nivel socioeconómico en Brasil y los Estados Unidos, las acciones que eran irrespetuosas o repugnantes también eran moralmente incorrectas (universalmente incorrectas e invariables), incluso cuando los encuestados declararon específicamente que nadie había resultado perjudicado por la acción.

Shweder (1990; Shweder, *et al*. 1997) ofreció más tarde una útil sistematización de la amplitud y variación del dominio moral. A partir del análisis por categoría del discurso moral que fue proporcionado por sus encuestados de Oriya, y de su propia lectura de la literatura antropológica, Shweder propuso que el discurso moral alrededor del mundo generalmente se basa en una o más de tres 'éticas': autonomía, comunidad y divinidad. Cada ética es un conjunto de afirmaciones morales interrelacionadas que funcionan para proteger a una entidad diferente. La 'ética de la autonomía' funciona para proteger a las personas, utilizando conceptos como daño y sufrimiento, derechos y justicia, libertad y autonomía. Este es el dominio moral, tal como lo define Turiel.

Sin embargo, en la mayoría de las culturas la gente cree que existen cosas aparte de los individuos que deben ser protegidas. La función de la 'ética de la comunidad' es proteger grupos, instituciones y otras entidades colectivas utilizando conceptos como deber, respeto, honor, lealtad y tradición. La función de la ética de la divinidad es proteger y glorificar a Dios, particularmente cuando Dios se manifiesta dentro de cada persona. Esta ética involucra conceptos morales como pureza, piedad, castidad y otras formas de autocontrol que ayudan a las personas a vivir de una manera más divina y menos carnal. Cuando se hacen comparaciones empíricas del discurso moral entre grupos más y menos occidentalizados (Haidt *et al*. 1993; Jensen 1998), o entre grupos más o menos políticamente conservadores y religiosos en los Estados Unidos (Haidt y Graham 2007, 2009; Haidt y Hersh 2001; Jensen 1997), se encuentra generalmente que, mientras los occidentales liberales seculares bien instruidos limitan en gran medida su discurso moral a la ética de la autonomía, otros grupos hacen uso de un conjunto más amplio de conceptos, basándose en gran medida en la ética de la comunidad y, a menudo (aunque no tan generalizadamente), en la ética de la divinidad.

Esta concepción más amplia de la moral plantea dos desafíos para los teóricos del innatismo. En primer lugar, deben explicar cómo el conocimiento o las respuestas a este conjunto completo de cuestiones morales, no solo el daño, los derechos y la justicia, son innatos. En segundo lugar, deben reconciliar su historia sobre lo innato con la variación obvia de las reglas y prácticas morales (y del dominio moral en sí) entre culturas.

CC BY-NC-ND

3. Cinco formas en que la moralidad podría ser innata

Teniendo en cuenta el tema de este volumen, nuestro objetivo aquí es explorar si un conjunto amplio y heterogéneo de conceptos y motivos morales — más allá del daño, los derechos y la justicia — podría reflejar la existencia de algún tipo de contenido mental dado de forma innata. Por tanto, lo mejor será que tengamos claro aquello que entendemos por 'innato'. La palabra ha sido utilizada de tantas formas por filósofos, biólogos y etólogos (Wimsatt 1999, enumera 13 significados distintos) que algunos académicos han abandonado la idea de que se trate de un concepto útil (v. g., Griffiths 2002). Ahora bien, en los escritos de Gary Marcus (2004), quien estudia las vías de desarrollo por las cuales los genes guían la construcción del cerebro, encontramos un enfoque simple y amigable. Marcus usa la metáfora de que los genes crean el *primer borrador* del cerebro, y la experiencia luego lo *edita*: "La naturaleza otorga al recién nacido un cerebro considerablemente complejo, pero uno que se comprende mejor como precableado, flexible y sujeto a cambios, en lugar de cableado, fijo e inmutable" (2004, 12). A su vez, Marcus explica que la edición (los cambios en el cerebro a medida que aprende y crece) se rige por procesos genéticos. Los genes no son solo patrones para producir proteínas, como se pensaba hace décadas. Más bien, una parte de cada gen está dedicada a procesos regulatorios: encender y apagar el gen en respuesta a diversas señales químicas. Marcus (2004, 40) explica que "'integrado' [*built-in*] no significa no maleable, sino que significa organizado antes de la experiencia" (en Samuels 2004, y en este volumen, se consideran muchos significados de lo innato y se llega a una conclusión similar). Nosotros adoptamos la visión de lo innato de Marcus, y tanto en esta sección como en la siguiente, tratamos de explicar las formas en que la moral humana puede estar "organizada antes de la experiencia". En las secciones 5 y 6 trataremos de explicar cómo la experiencia cultural y personal revisa el primer borrador (universal) durante el desarrollo infantil.

En lo que sigue describiremos cinco formas en que la moral podría ser innata. Comenzaremos con dos teorías del desarrollo moral —el constructivismo y el conexionismo— que son esencialmente teorías de *tabula rasa* en el sentido de que solo postulan procesos de aprendizaje generales innatos, no contenido moral innato (v. g., ideas, conocimiento). Sugeriremos que ambas teorías son parcialmente correctas como descripciones del proceso de edición, y que ambas pueden mejorarse al plantear al menos algún contenido que esté organizado antes de la experiencia. Luego describiremos tres enfoques que *sí* postulan un contenido moral innato específico del dominio. Creemos que estos cinco enfoques son útiles, particularmente cuando se distingue el primer borrador y los procesos de edición. En el apartado 4 propondremos nuestras propias hipótesis sobre aquello que está inscrito en el primer borrador de juicio moral, y en los apartados 5 y 6 describiremos cómo consideramos que está editado.

CC BY-NC-ND

3.1 Constructivismo piagetiano

Jean Piaget (1965/1932) se arrodilló y jugó a las canicas para estudiar el proceso por el cual los niños llegan a comprender las reglas. En contraste con sus contemporáneos que enfatizaron la internalización pasiva del niño de la moral del padre (Freud) o de la sociedad (Durkheim), Piaget mostró que los niños son participantes activos en su propio desarrollo. El desarrollo en cualquier dominio era, para Piaget, el producto de la interacción continua del niño con su entorno, como resultado de lo cual construía y reconstruía continuamente una comprensión progresivamente más adecuada (es por ello que la teoría de Piaget a menudo se denomina 'constructivismo'). Lawrence Kohlberg, en sus famosas seis etapas, elaboró la secuencia de comprensiones progresivamente más adecuadas de la moral, y atribuyó al proceso de 'toma de roles' la fuerza impulsora del desarrollo moral.

Para Piaget y Kohlberg, solamente hay una palabra escrita en el primer borrador de la mente moral: empatía (o quizás una palabra compuesta: 'toma de perspectiva'). También hubo algunas palabras escritas en otros capítulos (no morales) del primer borrador: me gusta y no me gusta. En la medida en que le gusten algunas cosas (como el placer, los dulces o la amistad) y no le gusten otras (como el dolor, la frustración o el rechazo), y que se se vuelva cada vez más hábil para tomar la perspectiva de los demás durante la etapa operativa concreta, el niño puede sentir por sí mismo (con empatía) que las acciones que dañan a otros son malas, mientras que las acciones que hacen felices a los demás son buenas. De esta manera, los niños llegan a comprender el valor de diferentes tipos de reglas y a apreciar que las reglas y las prácticas sociales deben justificarse por referencia a otra cosa. Para Kohlberg, esa otra cosa es la tradición, la autoridad y la sociedad para los razonadores morales 'convencionales', pero es la justicia para los razonadores morales más avanzados. El conocimiento de la justicia no es innato: es el logro culminante del proceso de edición, y el proceso de edición es un constructivismo impulsado por la experiencia de la toma de roles.

Este enfoque es elegante al explicar cuánto se puede derivar de tan poco conocimiento innato, y es razonable si acaso se considera que el dominio moral está restringido a cuestiones de daño, derechos y justicia. Sin embargo, si se comprende que conceptos como la obediencia, el respeto, el honor, la castidad, la templanza y el sacrilegio son conceptos efectivamente morales que precisan ser explicados (en lugar de ser superados por el niño en su camino hacia la autonomía moral y el razonamiento de la justicia perfecta), entonces no está claro cómo pueden derivarse de la empatía y de la adopción de roles, ¿por qué no postular que el primer borrador de la mente moral tiene varias palabras escritas, además de la empatía, y que Piaget y Kohlberg tienen razón respecto de que el constructivismo es parte del proceso de edición?

CC BY-NC-ND

3.2 Conexionismo

Un segundo enfoque que se centra en el proceso de edición es el paradigma conexionista propuesto por Paul Churchland (1996; 1998). Churchland parte del hecho manifiesto de que, además del entorno físico, los seres humanos nacen y viven en un mundo sociomoral sumamente denso y complejo. El problema para estas criaturas es aprender a navegar con éxito, desarrollando representaciones adecuadas de ese medioambiente y emparejando esas representaciones con respuestas conductuales apropiadas. No existe una facultad especial para realizar esta tarea. Se logra de la misma manera en que las personas aprenden a representar y vivir en sus entornos físicos: a través del ajuste gradual de la experiencia por parte de una mente producida por un cerebro, comprendido como una red neuronal.

El enfoque conexionista del funcionamiento moral de Churchland, es esencialmente una explicación del aprendizaje moral. Para Churchland, el desarrollo moral no significa, como significaba para Kohlberg, la formulación gradual de principios morales universales y abstractos. Más bien, se trata de "generar lentamente una jerarquía de prototipos morales, presumiblemente a partir de un número sustancial de *ejemplos* relevantes de los tipos morales en cuestión" (Churchland 1996, 102). La explicación de Churchland sobre la cognición y la competencia moral es estrechamente congruente con la visión moderna de la cognición, y también es congruente con algunas de las concepciones más antiguas del desarrollo moral, en particular la de Aristóteles (Casebeer 2003). De hecho, el enfoque de Churchland es esencialmente una traducción de los términos de la ética de la virtud al lenguaje conexionista. Trata la competencia moral como un conjunto de habilidades que se cultivan gradualmente, mediante la práctica, y que reciben la ayuda del énfasis de los adultos en las historias morales, las fábulas y los modelos a seguir.

El enfoque de Churchland es persuasivo sobre cómo ocurre el proceso de edición, pero guarda silencio sobre el contenido del primer borrador. De hecho, invita a la inferencia de que no hay un primer borrador más allá de cierto interés innato por las personas y los eventos sociales. Vemos esto como un defecto fácilmente corregible, ya que implica equipotencialidad en el aprendizaje moral: los niños podrían aprender fácilmente a navegar y valorar cualquier mundo moral artificial que los adultos crean para ellos, como el de las comunas populares[5] y los kibbutzim que intentaron, sin éxito, superar las preferencias de las personas por compartir bienes materiales con sus parientes cercanos y su aversión a procrear con las personas con las que se criaron. Así como los niños ingresan al mundo con algunas configuraciones iniciales en el dominio de la comida (un gusto por lo dulce, una aversión por lo amargo) que luego se amplían mediante el aprendizaje cultural, parece probable que los niños ingresen al mundo con algunas configuraciones iniciales en el dominio social (gusto por la justicia, disgusto por el daño) que luego se amplían con el aprendizaje cultural. Es poco violento para la teoría de Churchland proponer que el estado inicial del sistema, por ejemplo, la importancia originaria

[5] N. del T.: el autor se refiere a las 'comunas populares' [*people's commune*] en China que existieron desde 1958 a 1983.

CC BY-NC-ND

de ciertas conexiones sinápticas particulares, está "organizado antes de la experiencia" y luego es editado por la experiencia en la forma en que Churchland describe (véase la crítica de Clark (2000) a Churchland para un punto similar).

3.3 Modelos relacionales

Un tercer enfoque del innatismo es la teoría de los 'modelos relacionales' de Alan Fiske (1991; 1992). La teoría de Fiske es elegante y parsimoniosa, fue diseñada para explicar la similitud y la diversidad intercultural en una amplia gama de dominios culturales. Fiske (2004, 3) ofrece el siguiente resumen:

> La teoría de los modelos relacionales es simple: las personas se relacionan entre sí de solo cuatro formas. La interacción se puede estructurar con respecto a (1) lo que las personas tienen en común, (2) diferencias ordenadas, (3) desequilibrios aditivos o (4) proporciones. Cuando las personas se enfocan en lo que tienen en común, están usando un modelo que llamamos Comunidad[6]. Cuando las personas construyen algún aspecto de una interacción en términos de diferencias ordenadas, el modelo es Autoridad. Cuando las personas atienden a los desequilibrios aditivos, están enmarcando la interacción en términos del modelo de Igualdad. Cuando coordinan sus acciones según proporciones o tarifas, el modelo es Proporcionalidad. (...) En resumen, cuatro estructuras relacionales innatas y abiertas, completadas por complementos congruentes que son transmitidos socialmente, y que estructuran la mayor parte de la acción, el pensamiento y la motivación sociales. Esta es la teoría.

Fiske sugiere que los tres primeros modelos son innatos y se hallan con claridad en otros primates (Haslam 1997). No obstante, el modelo de Proporcionalidad (relaciones sociales basadas en proporciones) parece ser exclusivamente humano, y emerge más tarde en el desarrollo filogenético y ontogenético que el de Comunidad, Autoridad e Igualdad. El argumento a favor del carácter innato de la Proporcionalidad es, por tanto, mucho más débil, y Fiske sugiere que podría estar en proceso de convertirse en un mecanismo psicológico innato para las relaciones sociales. El autor, entonces, es muy claro respecto de que el primer borrador de la mente social contiene tres tendencias socioperceptivas primitivas e ineludibles: ver a las personas como grupos en los que todos son iguales, como ordenados por rango en una o más dimensiones, o como entidades cuya posición relativa debe mantenerse igual a pesar de los cambios periódicos fuera de equilibrio. Estos modelos innatos son utilizados en todas las culturas para estructurar el mundo social, aunque las culturas a menudo optan por usar diferentes combinaciones de modelos para así administrar cualquier relación determinada.

[6] N. del T.: Para la traducción de los nombres de los modelos de Fiske se optó por seguir la interpretación simplificada de Fernández, Cela-Conde y Fernández (2007). Una traducción más literal y descriptiva de los mismos podría ser Compartir Comunal (*Communal Sharing*), Rango de Autoridad (*Authority Ranking*), Precios del Mercado (*Market Pricing*) y Comparación Igualitaria (*Equality Matching*).

CC BY-NC-ND

Por ejemplo, la división del trabajo doméstico puede depender de la Autoridad en algunas culturas (por ejemplo, el marido es la autoridad y dicta quién hace qué), o de la Igualdad (ambos cónyuges se turnan en cada trabajo), o de la Comunidad (todos colaboran sin hacer un seguimiento de quién hace qué), o de la Proporcionalidad (a los trabajos se les asignan valores proporcionales a su dificultad, y los niños, o ayudantes contratados, reciben un pago por realizarlos).

La mayor parte de la teoría de los modelos relacionales de Fiske trata sobre el proceso de edición: cómo este primer borrador, con tres (o cuatro) marcos cognitivos simples y abiertos, se completa y ajusta durante el desarrollo infantil. No tenemos ninguna crítica a la teoría de Fiske, excepto que creemos que en el primer borrador hay un poco más, como explicaremos a continuación.

3.4 Modularidad masiva

El cuarto —y más ampleamente discutido— enfoque del innatismo es el concepto de modularidad. Es posible distinguir entre teorías *minimalistas* y *maximalistas* de la modularidad. Desde el punto de vista minimalista (Fodor 1983), un número muy pequeño de funciones mentales, que tienen que ver principalmente con la percepción sensorial y el lenguaje, son modulares, en el sentido de que son mecanismos computacionales innatos, rápidos, informacionalmente encapsulados y funcionalmente especializados. Fodor ofreció una definición específica de aquello que se precisa para ser un módulo. Luego afirmó que hay muy pocos módulos en la mente y que no existe un módulo para manejar tareas de orden superior, como el razonamiento o el juicio moral.

Según la visión maximalista propuesta por los psicólogos evolucionistas (v. g., Barkow, Cosmides y Tooby 1992, y este volumen; Buss 2004; Pinker 1997), gran parte de la mente está compuesta de módulos, cientos o miles de ellos, diseñados por la selección natural para resolver problemas específicos que eran recurrentes en el medioambiente ancestral. Según Tooby, Cosmides y Barrett (2005), muchos de estos módulos juegan un papel importante en nuestra vida moral. Argumentan que el estudio de la valoración, incluso más que otros tipos de cognición, revela cuán crucial es postular no solo procesos de aprendizaje innatos, sino también *contenido* mental innato. Los niños nacen con una preferencia (valoración) por la dulzura y un rechazo a la amargura. Cualquier padre sabe que la preferencia por los caramelos en lugar de por el brócoli no se aprende mediante la 'socialización' y no se puede deshacer con modelos a seguir, amenazas o recompensas. Tooby *et al*. sugieren que lo mismo es cierto para la valoración en todos los dominios:

> El contenido propio introducido por la arquitectura constituye una forma de conocimiento: la arquitectura debe saber (en cierto sentido) que los niños vivos son mejores que los niños muertos, que la aprobación social es mejor que la desaprobación, que la sal y lo dulce es mejor que lo agrio o putrefacto, que el sexo con tu madre o tu padre

CC BY-NC-ND

> deben ser evitados, que ayudar a tus hermanos es (dentro de ciertos intercambios) mejor que ayudar a los hongos, que tu pareja copulando con tu rival sexual es peor que su fidelidad, que las arañas en tu mejilla son peores que en el jardín, que la comprensión es mejor que la confusión, que el dominio de habilidades es mejor que la ineptitud, y así sucesivamente. (2005, 317)

En este pasaje, Tooby *et al.* reúnen muchos tipos de valoración bajo la rúbrica de lo que en otro lado denominan 'principios motivacionales', y no todos son relevantes para nuestro propósito, que es analizar las formas en las que específicamente el juicio *moral* podría tener una base innata. La sal, la dulzura y las arañas, por ejemplo, si bien son claramente objetos de gustos y preferencias, parecen ser diferentes en especie de las preferencias relacionadas con la comprensión, la fidelidad sexual y la ayuda, al menos por el hecho de que estas últimas parecen tener más contenido conceptual y menos perceptual. Tooby *et al.*, en lugar de una teoría de la moral, están interesados en desarrollar un inventario y una ciencia de la motivación *en general*, y si bien ambas están ciertamente relacionadas (y puede que, quizás, la primera resulte ser solo un caso especial de la segunda), el dominio moral es lo suficientemente distintivo como para que a su muy útil explicación le haga falta una pequeña modificación.

Estamos de acuerdo con Tooby *et al.* en que la valoración, tanto como para el comportamiento social como para la comida, es imposible de explicar si uno se niega a considerar la noción de que hay una estructura y un contenido innatos integrados en la mente. En sus palabras:

> (…) es necesario que haya un conjunto básico irreductible inicial, evolucionado y derivado de la arquitectura, de procedimientos de asignación de valor específicos en términos de contenido, o el sistema no podría ponerse en marcha. Ciertamente, el debate no puede ser sobre la existencia necesaria de este conjunto básico. La verdadera discusión es sobre qué tan grande debe ser el conjunto básico y cuál sería la descripción computacional adecuada de estos procedimientos de valoración y sus circuitos motivacionales asociados (2005, 317).

La noción de modularidad masiva es controvertida. Jerry Fodor, el autor original del concepto de módulos mentales, ha dicho que "evidentemente la tesis de la modularidad masiva *no es* cierta" (Fodor 2000, 23), y varios otros pensadores han seguido su ejemplo por diversas razones (ver, v. g., Buller 2005, Buller y Hardcastle 2000). Vemos dos dificultades principales al aplicar la tesis de la modularidad masiva a la moral.

La primera es una de las principales preocupaciones de Fodor, conocida como el 'problema de la flexibilidad'. La cognición humana de orden superior —y ciertamente la cognición moral— es bastante flexible. Las personas y las sociedades son bastante hábiles para invocar conceptos morales que se adaptan a sus propósitos, o torcer aquellos que no lo hacen para

CC BY-NC-ND

que posean formas más dóciles. Las reacciones que a menudo se afirma que están basadas en el conocimiento modularizado, como el miedo a las arañas, parecen tener una calidad de estímulo-respuesta de un nivel más bajo[7].

El segundo problema es el problema de la encapsulación: si bien muchos juicios morales cumplen con la mayoría de los criterios de modularidad de Fodor, incluida la especificidad de dominio y la velocidad, es inverosímil pensar que los juicios morales estén tan encapsulados informacionalmente como el tipo de fenómenos que suelen utilizarse para ilustrar la modularidad al nivel de la percepción. Por ejemplo, la ilusión de Müller-Lyer no se ve afectada por el conocimiento que uno tiene de las longitudes verdaderas de las líneas, pero los juicios morales se ven fácilmente afectados por el aprendizaje de nuevos hechos acerca de la situación en cuestión, por manipulaciones experimentales del estado de ánimo u otros factores que parecen ajenos al procesamiento de un módulo moral (v. g., Valdesolo y de Steno 2006). De modo que simpatizamos con la posibilidad de un conocimiento sustancial específico de dominio en el primer borrador de la mente moral, pero preferiríamos una versión de modularidad que pueda resolver estos dos problemas.

3.5 Modularidad 'abundante'

Varios teóricos (v. g., Carruthers[8]) han buscado un camino intermedio entre las concepciones totalmente no modulares de la mente y las teorías masivamente modulares. Estos pensadores hablan de "modularidad moderadamente masiva" o "modularidad en un grado significativo". Dudan de que no haya módulos conceptuales, pero también son escépticos de que la mente sea una navaja suiza repleta de herramientas que fueron completamente diseñadas hace mucho tiempo. Uno de los más importantes de estos modularistas moderados es el antropólogo Dan Sperber (1994; 2005). Como antropólogo, el objetivo de Sperber era explicar tanto la diversidad como la estabilidad de la cultura. La modularidad masiva con módulos fodoreanos es difícil (aunque no imposible) de reconciliar con la diversidad cultural de conceptos y patrones de comportamiento, así como también con el ritmo ocasionalmente rápido del cambio cultural. No obstante, por otro lado, los modelos de tabula casi rasa asumidos por

[7] Por supuesto, los módulos se pueden combinar creando sistemas que generan comportamientos flexibles, como se afirma respecto del denominado "órgano del lenguaje" (Pinker 1997). Pero debido a que no creemos que haya un solo órgano de la moral (ver Greene y Haidt 2002), buscamos módulos morales en un nivel más simple, como múltiples fuentes de intuición. Hauser (2006) afirma que existe un órgano de la moral, pero creemos que ha descrito simplemente un órgano que procesa daños.

[8] Disponible en http://www.philosophy.umd.edu/Faculty/pcarruthers/Moderate-modularity.htm.

CC BY-NC-ND

muchos antropólogos no pueden explicar ni las profundas y sorprendentes similitudes entre culturas (v. g., respecto de dioses, fantasmas y brujas; Boyer 2001), ni el grado de estabilidad intergeneracional que logran la mayoría de las culturas (Sperber y Hirschfeld 2004).

La solución de Sperber es una versión de la modularidad masiva, pero sus módulos son decididamente anti-fodoreanos: son muy variables (algunos cumplen con todos los criterios de Fodor, algunos solo cumplen con unos pocos), a menudo están anidados entre sí (al igual que el sistema digestivo es un módulo biológico que contiene muchos submódulos) y, lo que es más importante, la mayoría de los módulos de Sperber no son innatos, son generados durante el desarrollo por parte de un conjunto más pequeño de "módulos de aprendizaje" que son patrones innatos o "instintos de aprendizaje" (Sperber 2005, 57, citando a Marler 1991). Algunos de estos módulos innatos tienen un contenido perceptual específico incorporado. Por ejemplo, un módulo de aprendizaje de frutas "sabrá" que la fruta es dulce y luego solo generará submódulos de reconocimiento para las frutas posteriores (v. g., uno para manzanas, uno para plátanos), es decir, para objetos del medioambiente que cumplan con los criterios especificados previamente. Otros módulos de aprendizaje pueden ser más plenamente conceptuales. Por ejemplo, si hay un módulo de aprendizaje innato para la equidad, este genera una serie de módulos de detección de inequidad específicos de la cultura, como un 'detector de colarse en la fila' en culturas donde la gente hace cola, pero no en culturas donde no se hace, o un detector de 'división inequitativa de alimentos' en culturas donde los niños esperan recibir porciones exactamente iguales que sus hermanos, pero no en culturas donde las porciones se reparten por edades. Debido a que Sperber imagina un conjunto básico de módulos innatos que generan una gran diversidad de otros módulos, usa el sugerente término de 'modularidad abundante'.

En este punto, es probable que cualquier lector que no sea un teórico de la modularidad piense que nos hemos unido a Sperber mediante un salto desde un acantilado hacia una tierra donde todo objeto e individuo se denomina 'módulo'. Permítanos explicar por qué nos intrigan las ideas de Sperber. Nuestro objetivo es comprender el primer borrador y los procesos de revisión creados por la mente moral. Nuestra investigación empírica se centra en la intuición moral y el asombro moral (para Haidt) y en la cultura y la virtud (para Joseph). Ambos hemos descubierto que el juicio moral como la aplicación general de dominio de reglas y principios a casos específicos, como si el juicio moral fuera producto del razonamiento moral en el sentido kohlbergiano, no es una buena descripción. Más bien, cuando las personas son entrevistadas sobre situaciones en las que se viola un tabú (como el incesto consensuado entre hermanos o el canibalismo inofensivo), responden muy rápidamente y sus respuestas muestran lo que parece ser una especie de encapsulación similar a la de Müller-Lyer: mediante un interrogatorio, las personas a veces pueden ser empujadas a decir "No sé por qué, no puedo explicarlo, solo sé que está mal" (Haidt 2001; Haidt y Bjorklund 2008). Hemos argumentado (Haidt y Joseph 2004) que la mente adulta está llena de intuiciones morales, que son como pequeños fragmentos de programación *input-output* que conectan la percepción de un patrón en el mundo social (a menudo una virtud o vicio) con una evaluación, y en muchos casos, una

CC BY-NC-ND

emoción moral específica (por ejemplo, ira, desprecio, admiración). Cuando la gente piensa, chismea y discute sobre cuestiones morales, el terreno de juego no es afectivamente plano ni abierto a cualquier tipo de razón. Es más como un campo minado o una máquina de *pinball* donde destellos de intuición afectivamente cargada rebotan alrededor de nuestra atención y nos empujan hacia conclusiones específicas (ver Damasio 1994, para una idea similar, y para descripciones de lo que le sucede al pensamiento moral cuando se eliminan estos destellos). Estas intuiciones no son módulos de Fodor, pero sí son modulares "en un grado significativo" (Sperber 1994). Son fragmentos de la estructura mental, rápidos y de dominio específico, que influyen fuertemente en el juicio moral (Haidt 2001) ¿De dónde vienen todas estas intuiciones?

Quizás todas las intuiciones son innatas y la gente simplemente aprende qué eventos, en su cultura, cuentan como actos dañinos o inequidad (por ejemplo, colarse en la fila). No obstante, el enfoque de Sperber nos permite explicar ciertos *gustos morales adquiridos*, de una manera muy similar a como se explican otros tipos de gustos y temores adquiridos. Las personas se sienten atraídas de forma innata por las frutas y la carne, pero existen mecanismos especiales de aprendizaje que pueden generar un disgusto nuevo y duradero hacia alimentos específicos, en particular hacia las carnes. Por ejemplo, en 1805, cuando los hombres de la expedición de Lewis y Clark, luego de meses de inanición en las montañas Bitterroot de Idaho, entraron en contacto por primera vez con la tribu Nez Perce que se encuentra a lo largo del río Columbia, se atiborraron de salmón y tubérculos, ambos alimentos nuevos para ellos. Muchos se enfermaron esa misma noche, probablemente por los tubérculos apenas digeribles, pero debido a los mecanismos innatos de aprendizaje que asocian las náuseas con la carne (es mucho más probable que sea la carne la que contenga contaminantes bacterianos y no las verduras), desarrollaron un disgusto hacia el salmón. El disgusto fue tan fuerte que en los días siguientes los hombres compraron perros a los lugareños para comérselos, porque esa era la única carne disponible además del salmón (Burns 1997) ¿Acaso fue este un ejemplo de aprendizaje de dominio general? Fodor tendría que decir que sí, pero Sperber diría que no: era la génesis de un nuevo módulo a partir de un módulo de aprendizaje innato. Nosotros estamos de acuerdo con Sperber, vemos esto como una nueva intuición (un presentimiento) generada por un proceso de aprendizaje innato que puede alterar radicalmente el valor de las cosas sobre la base de la experiencia, pero solo dentro de los límites relacionados con la adaptación evolutiva. La nueva intuición estaba parcialmente encapsulada: si Lewis y Clark hubieran convencido a su tripulación de que los tubérculos eran los verdaderos culpables, los hombres habrían sentido disgusto por el salmón de todos modos.

El desarrollo moral muestra algunas de estas mismas características. Los niños llegan a reconocer gradualmente un gran conjunto de 'disparadores' a los que luego reaccionan rápida, automática y emocionalmente. Por ejemplo, en las últimas décadas, los estadounidenses han sintonizado mucho con el tema del abuso sexual de niños, tanto que les horrorizan patrones sociales que son bastante normales en otras partes del mundo, como el que los niños, durante la niñez media, duerman en la misma cama con el padre del sexo opuesto (Shweder,

CC BY-NC-ND

Balle-Jensen y Goldstein 1995), o besar los genitales de los bebés varones como expresión de afecto. Explicarles a los estadounidenses que los involucrados no creen que estas prácticas tengan algo que ver con la sexualidad no eliminará el disgusto y la condena, o la carga de abuso infantil ¿Existe un detector innato de abuso sexual? Probablemente no. Ahora bien, como explicaremos a continuación, sí creemos que hay algo innato —algo "organizado antes de la experiencia"— que hace que la actividad sexual y la protección de los niños sean dominios preparados evolutivamente para la preocupación moral. Otros ejemplos serían la velocidad, la facilidad y la pasión con la que la 'derecha religiosa' estadounidense ve el pecado, la tentación y el sacrilegio, o con la que la izquierda política estadounidense ve el racismo, la opresión y la victimización. Ya sea que estas reacciones morales se vean o no como manifestaciones de módulos adquiridos (abundantes) o simplemente como experiencias dependientes de la cultura, son ejemplos de intuiciones morales: fragmentos de la estructura mental que conectan la percepción de patrones específicos en el mundo social con evaluaciones y emociones que no son totalmente controlables o revisables por la persona que las experimenta.

4. El primer borrador de la moral: los cinco fundamentos de la ética intuitiva

Durante mucho tiempo hemos estado buscando los fundamentos de las intuiciones éticas: aquellas bases psicológicas primitivas que son los ladrillos de construcción a partir de los cuales las culturas crean morales únicas pero limitadas en sus variaciones. Recientemente (Haidt y Joseph 2004), examinamos varias teorías (incluidas las de Shweder y Fiske) respecto de la amplitud de la moral humana y sus aspectos precursores en otros primates (por ejemplo, de Waal 1996). Intentamos identificar la gama completa de fenómenos a través de diferentes culturas, aquella que precisaría ser explicada por cualquier teoría adecuada de la moral humana. A su vez, identificamos cinco conjuntos de preocupaciones, cada uno vinculado a un desafío adaptativo y a una o más emociones morales, como los mejores candidatos de los fundamentos psicológicos de la moral humana. Los cinco fundamentos que identificamos son Daño/Cuidado, Equidad/Reciprocidad, Intragrupo/Lealtad, Autoridad/Respeto y Pureza/Santidad. Cada uno de estos cinco es un buen candidato para un módulo de aprendizaje al estilo de Sperber. Sin embargo, para aquellos lectores que no gustan de las teorías de la modularidad, pueden comprenderlas como disposiciones evolutivas (Seligman 1971) que vinculan ciertos patrones de valoración social con reacciones emocionales y motivacionales específicas. Lo que nos interesa enfatizar es que la mente moral está parcialmente estructurada antes de la experiencia, de modo que es probable que cinco (o más) clases de preocupaciones sociales adquieran una dimensión moral durante el desarrollo. Aquellos problemas sociales que no pueden relacionarse con alguna de estas bases son mucho más difíciles de enseñar o de inspirar preocupación.

La tabla 1 presenta nuestra teoría de manera concisa. La primera fila enumera cinco desafíos adaptativos de larga data que los mamíferos altamente sociales, como nuestros antepasados, enfrentaron durante millones de años, creando condiciones que favorecieron el éxito

CC BY-NC-ND

reproductivo de los individuos que pudieron resolver los problemas de manera más efectiva. Para cada desafío, la adaptación efectiva significó ser capaz de detectar ciertos patrones en el mundo social y responder a ellos con un perfil motivacional alterado. Sperber (1994) se refiere al conjunto de objetos para el cual un módulo fue 'diseñado', como el *dominio adecuado* para ese módulo. A su vez, contrasta el dominio adecuado con el *dominio real*, que es el conjunto de objetos que de hecho activan el módulo. Incluso cuando un animal vive en el entorno para el que está adaptado, el dominio real nunca coincide perfectamente con el dominio adecuado: el módulo produce falsos positivos y falsos negativos. En relación con una especie como la nuestra, que ha experimentado un rápido cambio ambiental y que manipula deliberadamente su medioambiente físico y social por múltiples razones, podemos estar bastante seguros de que el dominio real diverge del dominio adecuado en un grado sustancial: con seguridad nuestras intuiciones morales se involucrarán —o se suprimirán— de maneras que no contribuirán al éxito darwiniano de cada individuo.

	Daño/Cuidado	**Equidad/ Reciprocidad**	**Intragrupo/ Lealtad**	**Autoridad/ Respeto**	**Pureza/ Santidad**
Cambio adaptativo	Proteger y cuidar a los familiares jóvenes, vulnerables o heridos	Cosechar los beneficios de la cooperación diádica con personas que no son parientes	Aprovechar los beneficios de la cooperación grupal	Negociar la jerarquía, diferir de manera selectiva	Evitar microbios y parásitos
Dominio propio (disparadores adaptativos)	Sufrimiento, angustia o amenaza a los familiares	Trampa, cooperación, engaño	Amenaza o desafío al grupo	Signos de dominación y sumisión	Productos de desecho, personas enfermas
Dominio actual (el conjunto de todos los disparadores)	Crías de foca, personajes de dibujos animados	Fidelidad marital, máquinas expendedoras rotas	Equipos deportivos de los cuales se es hincha	Jefes, profesionales respetados	Ideas tabú (comunismo, racismo)
Emociones características	Compasión	Ira, gratitud, culpa	Orgullo de grupo, pertenencia; furia contra los traidores	Respeto, miedo	Asco
Virtudes relevantes [**y vicios**]	Cuidado, amabilidad, [crueldad]	equidad, justicia, honestidad, honradez [deshonestidad]	Lealtad, patriotismo, sacrificio personal[traición, cobardía]	Obediencia, deferencia [desobediencia, arrogancia]	Templanza, castidad, piedad, limpieza [lujuria, intemperancia]

Tabla 1. Los cinco fundamentos de la ética intuitiva.

CC BY-NC-ND

La forma de leer la tabla es de arriba abajo por columna. Por ejemplo, el fundamento Daño/Cuidado puede entenderse comenzando con el hecho de que los mamíferos, por definición, se enfrentan a la necesidad de cuidar a sus crías vulnerables y nada podría ser más central para el éxito evolutivo que mantenerlas con vida. Por lo tanto, es poco plausible que los mamíferos aprendan solamente a través de mecanismos de aprendizaje de dominio general, cómo reconocer el sufrimiento o la angustia en su descendencia. Más bien, muchos mamíferos tienen módulos innatos de detección de daños que fueron moldeados por la evolución para responder al dominio adecuado de los signos de sufrimiento en su propia descendencia. En la práctica real, este módulo (o conjunto de módulos) responde a muchas otras cosas además del sufrimiento y la angustia de los hijos propios. Para los humanos, el sufrimiento o daño a casi cualquier entidad infantil es parte del dominio real de este módulo (un cartel que muestre crías neoténicas de foca siendo apaleadas hasta la muerte por hombres grandes, representa un superestímulo deliberadamente ideado para este módulo). Estos módulos generalmente tienen como uno de sus resultados la emoción de la compasión: el individuo está motivado a actuar con el fin de proteger o aliviar el sufrimiento del niño. No sabemos si existe un único módulo de daño que posee desencadenantes tanto innatos como aprendidos (como sugiere Hauser 2006), o si el relato de la modularidad abundante es correcto en el argumento de que la mente humana está preparada de forma innata para generar una serie de módulos específicos relacionados con el daño. Sin embargo, si todas las personas tienen una sensibilidad emocional al daño, particularmente al daño efectuado a los débiles o vulnerables, y si todas poseen lenguaje, es probable que desarrollen un vocabulario que les permita hablar sobre sus reacciones emocionales. A su vez, es probable que tengan palabras sobre virtudes y vicios con las que alabar y condenar a las personas e instruir a sus hijos. Este discurso sobre las virtudes puede retroalimentarse para así afinar los límites y las aplicaciones de los módulos: las culturas pueden volverse expertas en percibir ciertos tipos de daño (v. g., abuso sexual o brujería).

En lo que sigue contaremos una historia similar para cada una de las otras cuatro columnas. El fundamento Equidad/Reciprocidad, por ejemplo, es solo una elaboración de la narración de Trivers (1971) sobre cómo pudo haber evolucionado un conjunto de emociones que ayuda a los organismos sociales a cosechar los beneficios del altruismo recíproco con parientes no familiares o lejanos. Junto con los atributos evolucionados a nivel individual de mayor interés y reacción emocional a signos de engaño y cooperación, viene un conjunto de productos culturales, como virtudes y vicios, palabras relacionadas con la equidad, mandatos religiosos sobre la reciprocidad, construcciones culturales como derechos e instituciones sociales relacionadas con la justicia.

La siguiente columna, del fundamento Intragrupo/Lealtad, organiza fenómenos relacionados con la muy estudiada tendencia humana a agruparse en tribus, pandillas y equipos que compiten con otras tribus, pandillas y equipos. Los experimentos de grupos mínimos de Tajfel *et al*. (1971) demuestran que las personas formarán tales grupos sobre la base de similitudes incluso triviales. Los grupos basados en sangre, religión o idioma compartidos son mucho

CC BY-NC-ND

más poderosos. Los conflictos por el territorio o los ataques de otros grupos parecen apelar con especial intensidad a las virtudes relacionadas con este fundamento, como la lealtad, el heroísmo y el autosacrificio por el bien común. Cuando estas virtudes intragrupales son prominentes, es probable que el grupo se muestre muy vigilante y punitivo hacia los traidores, los aprovechados y los holgazanes. El potencial destructivo de este módulo se exhibe diariamente en todo el mundo, incluidos episodios de genocidio y limpieza étnica. La moralización de la unificación durante la guerra es evidente también en el título de un libro reciente de un archiconservador estadounidense, indignado por la disidencia durante la guerra de Irak: *Treason: Liberal treachery from the cold war to the war on terrorism* (Coulter 2003).

La cuarta columna trata sobre los concomitantes psicológicos y sociales de la vida en las jerarquías de dominación. Muchos primates viven en tales jerarquías, y los patrones comunes de dominio y sumisión entre especies y culturas sugieren fuertemente que hay algo en la mente humana que se organizó antes de la experiencia, lo cual facilita que los humanos desarrollen un conjunto de emociones y comportamientos relacionados con la autoridad y el poder (ver Boehm 1999, para una discusión sobre cómo surgen las sociedades igualitarias a pesar de la predisposición humana a la vida jerárquica). Sin embargo, como Fiske señala repetidamente, la Autoridad es una calle de dos vías: los subordinados deben mostrar respeto y deferencia, pero los superiores deben protegerlos de amenazas externas y mantener el orden dentro del grupo. Este lado prosocial de la autoridad parece pasar desapercibido en muchas explicaciones psicológicas contemporáneas de la jerarquía que, como Shweder *et al.* (1997) señalan, ven todas las formas de desigualdad como formas de opresión. En las sociedades que valoran la autoridad, sin embargo, las normas y virtudes relacionadas gobiernan el comportamiento de los superiores (v. g., imparcialidad, magnanimidad, paternidad) y subordinados (v. g., respeto, deferencia).

Finalmente, el quinto fundamento, la pureza, es único en el sentido de que es el único para el cual el desafío adaptativo original no era social, sino nutritivo. La estrategia de alimentación omnívora de los seres humanos, combinada con nuestro tamaño grupal relativamente grande (en comparación con otros primates; Dunbar 1993) significa que durante mucho tiempo hemos estado expuestos a niveles muy altos de amenaza, particularmente por bacterias y parásitos que se propagan por contacto físico. Los humanos (pero no otros animales), por lo tanto, desarrollaron un conjunto de adaptaciones cognitivas y emocionales relacionadas con el disgusto, las cuales nos hacen ser cautelosos, pero al mismo tiempo flexibles sobre el tipo de cosas que comemos y sobre el historial de contacto para con las cosas que comemos (Rozin y Fallon 1987). Este sistema de evaluación y rechazo de alimentos se adaptó bien para la evaluación y rechazo social. De hecho, la mayoría, si no todas las sociedades humanas, utilizan parte del vocabulario y la lógica del disgusto físico en su vida moral (Haidt, *et al.* 1997; Rozin, Haidt y McCauley 2000). En algunas sociedades, la capacidad de rastrear el contagio y valorar la pureza parece contribuir a las ideas acerca de lo sagrado: sobre mantener los objetos religiosos separados de los objetos contaminantes y profanos, y sobre la superación

CC BY-NC-ND

de los deseos carnales y el tratamiento del cuerpo como un templo (véase Eliade 1957, sobre la sacralidad). Por tanto, este fundamento genera a menudo virtudes como la castidad y la templanza, y vicios como la lujuria y la intemperancia.

La pureza a menudo está profundamente moralizada, no solo como una preocupación por uno mismo, sino en forma de creencias y sentimientos sobre los grupos y el mundo en su conjunto. Esta es una fuente de lo que podría llamarse el 'lado oscuro' de las intuiciones de pureza y, de hecho, una preocupación (u obsesión) por la pureza a menudo se asocia con una violencia y opresión horribles, particularmente cuando se combina con intuiciones del fundamento Intragrupo. Por ejemplo, el holocausto, la 'limpieza' étnica y las leyes Jim Crow en el sur de Estados Unidos que mantuvieron los cuerpos y procesos corporales de los afroamericanos separados de los de los blancos.

Creemos que estos cinco conjuntos de problemas, sensibilidades y habilidades de percepción social son los mejores candidatos a la hora de comprender los cimientos de la ética intuitiva, por varias razones. En primer lugar, encontramos un sorprendente grado de similitud en la forma en que las culturas afrontan estos cinco desafíos adaptativos. Por ejemplo, en la lógica de los ritos de iniciación que crean un intragrupo fuerte, en las formas en que se marcan la jerarquía y la sumisión, y en las reglas de pureza y contaminación que tan a menudo regulan procesos biológicos como la menstruación, el nacimiento y la defecación. En segundo lugar, cuatro de los cinco fundamentos propuestos (todos menos la pureza) parecen involucrar "bloques psicológicos elementales" que están presentes en otros primates (de Waal 1996, aunque la reciprocidad todavía se debate; Hauser 2006), lo que nos da más confianza respecto de que hay algo de estos fundamentos que es "especificado antes de la experiencia". En tercer lugar, nuestros cinco fundamentos encajan perfectamente con las tres éticas de Shweder (los fundamentos de Daño y Equidad dan lugar al discurso de la ética de la autonomía; los fundamentos de Intragrupo y Autoridad apoyan la ética de la comunidad; y el fundamento de Pureza apoya la ética de la divinidad). En cuarto lugar, tres de nuestros fundamentos coinciden con los primeros tres modelos relacionales de Fiske (Equidad = Igualdad; Intragrupo = Comunidad; Autoridad = Autoridad). El hecho de que nuestros cinco fundamentos no coincidan con los cuatro modelos de Fiske, se debe a que el Daño y la Pureza no son principalmente modos de relación interpersonal. Los incluimos porque son fuentes importantes y probablemente innatas de valoración moral humana. Por otro lado, no incluimos el modelo de Proporcionalidad porque creemos que no resulta tan claro que sea innato. Fácilmente podríamos estar equivocados al excluirlo. A su vez, no afirmamos que solamente hay cinco fundamentos. Probablemente haya muchos más, pero creemos que los cinco que hemos identificado son los más importantes para explicar la moral humana y la diversidad moral.

5. El proceso de edición: desarrollando virtudes

Los cinco fundamentos son, por proponer una analogía, las 'papilas gustativas' innatas del sentido moral. La lengua humana tiene cinco tipos de receptores, cada uno de los cuales

CC BY-NC-ND

traduce un patrón químico de una sustancia a una experiencia afectiva que es positiva (para dulce, salado y glutamato) o negativa (para amargo y, más allá de cierto nivel de intensidad, para agrio). Estas papilas gustativas nos dicen algo sobre cómo vivían nuestros antepasados: comían fruta y carne, y tenían una variedad de herramientas de percepción en sus lenguas (y narices y ojos) que en sus cerebros se combinaban con herramientas conceptuales con el fin de guiarlos hacia la fruta y la carne. De manera similar, los cinco fundamentos sugieren algunas cuestiones sobre cómo vivían nuestros antepasados: eran criaturas ultrasociales (Richerson y Boyd 1998), en última instancia sintonizados para 1) criar hijos y ayudar a sus parientes, 2) cooperar selectivamente con individuos con los cuales no compartían parentesco mientras permanecían atentos a los tramposos, 3) formar intragrupos fuertes por la competencia extragrupal, 4) organizarse jerárquicamente, y 5) atender a los estados físicos de los demás, así como también alterar las interacciones y contactos de acuerdo con ello. Mientras que las papilas gustativas de la lengua recopilan información perceptiva (sobre azúcares, ácidos, etc.), las papilas gustativas del sentido moral responden a patrones conceptuales más abstractos (como el hacer trampa, la falta de respeto o la traición). No obstante, en ambos casos, el resultado es una experiencia con valencia afectiva (me gusta, no me gusta), la cual guía las decisiones posteriores respecto de si acercarse o evitar el objeto/agente en cuestión.

Por supuesto, hay mucho más en el juicio moral (y en la selección de alimentos) que la acción de cinco 'papilas gustativas'. El funcionamiento moral maduro no consiste solamente, ni siquiera principalmente, en simples reacciones afectivas o intuitivas a los estímulos sociales. El disgusto que se siente hacia las heces de los perros, o incluso hacia un acto de relación homosexual, no es en sí mismo un juicio moral. El desarrollo moral también se caracteriza por la adquisición y el uso de una amplia variedad de conceptos morales. Algunos de estos son categorías de acción: mentiras, traiciones, favores, etc. Otros son categorías de personas, o más específicamente, categorías de características de personas, las cuales incluyen rasgos con valencia positiva[9] como la bondad, la lealtad y la honradez, y rasgos con valencia negativa como la crueldad, la deshonestidad y la cobardía.

Estos rasgos, virtudes y vicios, están comenzando a resurgir en la psicología moral empírica después de un largo período de exilio, ocasionado en parte por las críticas de Lawrence

[9] La palabra 'rasgo' está cargada de significado en psicología. En particular, es el foco de un acalorado debate entre psicólogos sociales y psicólogos de la personalidad. Algunos han colocado los rasgos en el centro del estudio de la personalidad, mientras que otros, por diversas razones, se muestran escépticos o desdeñan el concepto en sí mismo (Mischel 1968; Ross y Nisbett 1991). Este debate ha penetrado en las discusiones filosóficas y psicológicas de la moral con 'situacionistas' como John Doris (1998) y Gilbert Harman (1999), quienes interpretan las virtudes como rasgos pero luego descartan su existencia, y los teóricos de la virtud (Sreenivasan 2002; Merritt 2000), quienes defienden algunas versiones de la teoría de la virtud frente a la crítica situacionista. Creemos que son los teóricos de la virtud los que tienen razón. Las virtudes, tal como las interpretamos, son *habilidades* o *capacidades* altamente vinculadas con situaciones específicas, más que disposiciones generales de comportamiento. Esta forma de ver las virtudes obvia la acusación básica de los situacionistas y es consistente con la crítica original y vigente de Mischel hacia las teorías de los rasgos (Cervone y Shoda 1999), y con el trabajo reciente en neurociencia cognitiva (v. g., Casebeer 2003; Churchland 1998).

CC BY-NC-ND

Kohlberg y otros teóricos. Anteriormente hemos discutido el papel que pueden jugar las virtudes en una teoría integral del funcionamiento moral y el desarrollo moral (Haidt y Joseph 2004). A continuación, repetiremos nuestros puntos principales, como preludio de una argumentación sobre un aspecto de la teoría de la virtud y de nuestra teoría en particular, a saber, la narrativa.

En primer lugar, ¿qué es una virtud? Existen muchos puntos de vista, pero la mayoría de los teóricos de la virtud estarían de acuerdo al menos en que son características, moralmente dignas de alabanza, de una persona. Por lo tanto, las virtudes son rasgos, siempre y cuando uno no piense en los rasgos como tendencias globales a actuar de una manera particular (v. g., honesto, valiente) en circunstancias muy diversas. Más bien, pensamos en los rasgos como lo hizo John Dewey: como patrones dinámicos de percepción, emoción, juicio y acción (Dewey 1922; ver también Churchland 1998). Las virtudes son habilidades sociales. Poseer una virtud es haber extendido y refinado las habilidades propias de percibir información moralmente relevante de modo que uno pueda responder plenamente al contexto sociomoral local. Ser amable, por ejemplo, es tener sensibilidad perceptiva hacia ciertas características de las situaciones, incluidas las que tienen que ver con el bienestar de los demás, y cierta propensión a que las motivaciones propias se moldeen y afecten adecuadamente. Ser valiente es tener un tipo de sensibilidad diferente. Ser paciente, otro más[10].

Uno de los principios cruciales de la teoría de la virtud es que las virtudes se adquieren de forma inductiva, a través de la exposición (a veces con un esfuerzo por copiar), muchos ejemplos de la virtud en la práctica. Cada uno de estos ejemplos contiene información sobre una serie de aspectos de la situación determinada, incluidas las motivaciones de los protagonistas, su estado (sufrimiento, discapacitados, hostiles, ricos, etc.), la categorización de la situación y la evaluación del resultado ofrecida por otros más experimentados. Solamente con el tiempo el aprendiz moral reconocerá a qué información es importante prestarle atención y retenerla, y cuál puede ser ignorada sin inconvenientes.

Los filósofos y los científicos cognitivos han argumentado recientemente, con respecto tanto a la moralidad como a la cognición en general, que este tipo de aprendizaje no puede ser reemplazado por un aprendizaje de arriba hacia abajo, como sucede con la aceptación de una regla o principio y la deducción de respuestas específicas. Curiosamente, este aspecto de la teoría de la virtud evidencia que Aristóteles fue un precursor de la aplicación del enfoque conexionista de la moral que hemos descrito anteriormente (ver May, Friedman y Clark 1997). En este modelo, la mente, como el cerebro mismo, es una red que se sintoniza gradualmente con la experiencia. Con entrenamiento, la mente logra hacer un trabajo progresivamente mejor, reconociendo patrones importantes de entrada y respondiendo con el patrón apropiado de salida.

[10] Para una exposición clásica de la interpretación de las virtudes como sensibilidades o capacidades perceptivas, véase McDowell (1979).

CC BY-NC-ND

Entonces, para aquellos que enfatizan la importancia de las virtudes en el funcionamiento moral, la madurez moral consiste en lograr una sintonía comprensiva con el mundo, un conjunto de sensibilidades altamente sofisticadas corporeizadas en las virtudes individuales. El razonamiento y la deliberación también juegan un papel importante en esta concepción: parte de ser una persona virtuosa es poder razonar correctamente sobre situaciones difíciles o problemáticas. Sin embargo, la teoría de la virtud implica un alejamiento de las teorías de la moral que ven la deliberación como la actividad psicológica moral básica.

La teoría de la virtud postula un tipo particular de organización de la competencia moral, una en la que la percepción, la motivación, la acción y el razonamiento corresponden a exigencias que se le imponen a una persona por las características particulares de cierta situación. Resulta claro que la objetividad de estas exigencias y la relevancia moral de las características de una situación están en cierta medida dictadas por la cultura, por los conceptos morales, las estructuras sociales y las narrativas que son corrientes en el contexto social inmediato. No obstante, esto no significa que el contenido o la estructura de una virtud sean culturalmente relativos por completo. Tal como señaló Aristóteles, y como han elaborado los especialistas en ética de la virtud (Nussbaum 1993), el hecho de que cierta característica de la personalidad sea una *virtud*, y no simplemente una regularidad conductual, consiste mayormente en su funcionar bien en una 'esfera de existencia' específica. A su vez, lo que Aristóteles y Nussbaum quieren decir con 'esferas de existencia' es similar a lo que los biólogos evolutivos reconocerían como desafíos adaptativos persistentes y otros tipos de restricciones ambientales. Por lo tanto, las virtudes son bien acogidas por una teoría científica del funcionamiento moral basada en la psicología evolucionista[11] y la psicología cultural.

La intersección de la teoría de la virtud, la ciencia cognitiva y el estudio empírico de la moral recién comienza a explorarse seriamente, y aún quedan muchas preguntas abiertas. Hemos discutido las explicaciones conexionistas de la cognición, así como también su extensión a la percepción social y la cognición social, particularmente porque los teóricos conexionistas (Casebeer, Churchland) se han interesado notablemente por la cognición social y moral. Ahora bien, es muy probable que alguna otra explicación de la cognición sea igualmente adecuada para describir y explicar la moralidad.

Nuestro compromiso con la teoría de la virtud en particular, sin embargo, es más firme. En nuestra opinión, esta teoría reúne aquello que es verdadero y útil en los cinco enfoques que esbozamos en la sección 3. El niño es, de hecho, un participante activo en su propio desarrollo. El conocimiento y las habilidades morales no se 'descargan' simplemente en la mente del niño, como lo dirían las teorías de socialización de tabula rasa. Piaget y Kohlberg

[11] N. del T.: Aunque sería más apropiado denominarla 'psicología evolutiva', se optó por el término 'evolucionista' para evitar confusiones para con la teoría psicológica de Jean Piaget (de hecho, gran parte de las traducciones de textos pertenecientes a esta vía de investigación en Psicología, suelen preferir el término 'evolucionista' por la misma razón). A su vez, si bien más adelante Haidt alude a la perspectiva piagetiana, en este caso se refiere a la psicología desarrollada, por ejemplo, por Tooby, Cosmides y Barrett (2005), citados en este capítulo.

CC BY-NC-ND

tienen razón en que hay un elemento sustancial de autoconstrucción en el desarrollo moral. Sin embargo, lo que se aprende es mejor descrito como aquellas habilidades de percepción social y reacción argumentadas por los conexionistas y los teóricos de la virtud. La mayoría de estas habilidades tratan sobre cómo interactuar con otras personas (cómo llenar los tres —o cuatro— modelos dados de manera innata para las relaciones sociales, descritos por Fiske). Sin embargo, parte de este conocimiento no consiste en relaciones *per se*. Hay mucho más que es innato, particularmente cuando miramos los orígenes de la valoración, como lo describen Tooby *et al.* (2005). Además de estar "organizada antes de la experiencia" para los primeros tres modelos de Fiske (que involucran grupos internos, autoridad y reciprocidad), la mente también está preparada de manera innata para percibir y preocuparse por el daño desde una edad muy temprana (Zahn-Waxler *et al.* 1979), y también por el disgusto, la pureza y la contaminación (desde una edad posterior, quizás recién a los 8 años; Rozin, Haidt y McCauley 2000). Es posible que Turiel haya tenido razón al centrarse en el daño, así como también en la capacidad del niño para comprender y disgustar del sufrimiento, como la intuición más importante del desarrollo moral temprano. Sin embargo, creemos que se equivocó al sugerir que los niños derivan el resto de sus otros conceptos morales autoconstruyéndolos a partir de esta única base.

En síntesis: la trayectoria de desarrollo característica en el dominio moral es un movimiento desde juicios globales crudos y articulados, que utilizan un pequeño número de intuiciones morales innatas para percepciones, creencias, respuestas emocionales y juicios altamente sofisticados y diferenciados. Esto es coherente con la noción de Sperber de 'modularidad abundante': los miniprogramas intuitivos, similares a módulos y de dominio específico dan lugar, en el agente moral maduro, a un conjunto expansivo y flexible de módulos morales que son más poderosos y sutiles que los módulos innatos, a saber, aquellos que componen los cinco fundamentos de la ética intuitiva que venimos discutiendo. El enfoque de Sperber sugiere que las virtudes no son innatas en sí mismas, sino que se adquieren a través de un proceso generativo en el que las capacidades de dominio específico de los módulos que componen las cinco bases se multiplican, amplían y refinan.

6. El proceso de edición: narrativas de aprendizaje

Constructivistas como Piaget y Kohlberg dirigieron nuestra atención a las formas en que los niños crean de manera activa una comprensión cada vez más matizada de las cuestiones morales. Kohlberg estudió los aspectos más explícitos, discursivos y deliberativos del funcionamiento moral. El desarrollo en etapas se midió a través del análisis del razonamiento verbal utilizado para justificar las respuestas a dilemas morales hipotéticos. En la tradición cognitivo-evolutiva, el pensamiento moral se consideraba similar al pensamiento lógico. Piaget (1965/1932, 398) afirmó explícitamente que "la lógica es la moral del pensamiento, así como la moral es la lógica de la acción".

CC BY-NC-ND

Es cierto que los niños reflexionan sobre cuestiones morales, particularmente cuando discuten con otros, y resulta plausible plantear un espacio de trabajo de dominio general donde se lleva a cabo el pensamiento moral (así como otros tipos de pensamiento). No obstante, ¿todo pensamiento moral verbal consciente debe comprenderse como pensamiento lógico?, ¿piensan realmente los niños tanto en los principios morales como en las formas en que se aplican o no en una situación determinada? Hay otro tipo de pensamiento, un tipo diferente de racionalidad que parece jugar un papel crucial en el pensamiento y desarrollo moral. Jerome Bruner (1990) distinguió entre el modo *narrativo* de la cognición y el modo paradigmático o *lógico-científico*:

> Hay dos modos de funcionamiento cognitivo, dos modos de pensamiento, cada uno de los cuales proporciona formas distintivas de ordenar la experiencia, de construir la realidad…. Una buena historia y un argumento bien formado son dos clases naturales diferentes. Ambos pueden usarse como un medio para convencer a un otro. Sin embargo, de qué convencen es fundamentalmente diferente: los argumentos convencen de su verdad, las historias de su realismo. El primero verifica mediante una eventual apelación a procedimientos que permiten establecer pruebas formales y empíricas. El segundo se funda no en la verdad sino la verosimilitud. (Bruner 1990, 11)

Bruner observa que sabemos mucho más sobre el modo paradigmático de pensamiento porque la psicología cognitiva ha concentrado su atención en él, mientras que, por otra parte, el modo narrativo ha sido ignorado. Más importante aún para los fines de este ensayo, Bruner también señala que cada modo de pensamiento se fundamenta en sus propias 'prótesis': ayudas al pensamiento proporcionadas por una cultura. Para el modo paradigmático, las prótesis incluyen lógica, matemáticas y ciencias. Para el modo narrativo, los dispositivos protésicos más comunes son los textos. La función de los textos, entre muchas otras, es almacenar significados culturales y, tanto por su contenido como por su estructura, ayudar a orientar el pensamiento de las personas. Creemos que el pensamiento, la argumentación y la reflexión moral (fuera de los departamentos de filosofía, al menos) se describen mucho mejor como una especie de pensamiento narrativo que como pensamiento paradigmático.

Hay muchos tipos diferentes de narrativas, de niveles de complejidad, y, como resultado, la narratividad[12] da forma al funcionamiento moral y al desarrollo moral en varios niveles de organización. Algunas de las narrativas morales más poderosas son las más simples. Para los occidentales, parábolas como las que se encuentran en las enseñanzas de Jesús del Nuevo Testamento son ejemplos familiares. Otras religiones tienen herramientas similares, por ejemplo, el *hadith* o los dichos y hechos del profeta Mahoma en el Islam. Shweder y Much (1991) encontraron que las narrativas se invocan comúnmente en las culturas hindúes como un modo de argumento moral. En sus entrevistas sobre dilemas morales con informantes indios, encontraron que las preguntas sobre lo correcto o incorrecto de un acto en particular suscitaron, a menudo, una respuesta que comenzaba con un "Déjame contarte una historia sobre..."

[12] Para una discusión de este concepto, véase Carruthers (1991).

CC BY-NC-ND

(por ejemplo, robar). De manera reveladora, cuando Shweder envió textos de entrevistas de este tipo a Kohlberg para que los puntuara, Kohlberg informó que gran parte del material de la entrevista no podía ser codificado desde su sistema.

Más recientemente, el sociólogo Christian Smith (2003, 64) ha observado que somos "animales que hacen historias, pero también animales hechos por nuestras historias". Smith describe una variedad de narrativas de alto orden, a menudo inconscientes, que organizan la identidad y el juicio moral tanto a nivel individual como grupal. Por ejemplo, señala que los estadounidenses y los "musulmanes militantes" interpretan los ataques del 11 de septiembre a la luz de metanarrativas muy diferentes: los estadounidenses ven las cosas a través de lo que Smith llama la "narrativa del experimento americano", según la cual los estadounidenses huyeron de la opresión del viejo mundo y desde entonces han sido un faro brillante de libertad y esperanza; por otro lado, la narrativa del "resurgimiento islámico militante" ofrece una visión radicalmente diferente, en la que Estados Unidos ha sido durante mucho tiempo un matón y un hipócrita. Hay otras narrativas, cada una de las cuales Smith explica casi como una receta. Entre ellas se encuentran la narrativa de la "prosperidad capitalista", la narrativa del "socialismo progresivo", la narrativa del "romántico expresivo" y la narrativa de la "Ilustración científica".

El aporte de Smith es especialmente útil para hacer explícitas las narrativas que motivan y guían a los sociólogos estadounidenses y otros académicos. Por ejemplo, la narrativa del "progreso liberal" cuenta la historia de cómo:

> Érase una vez, la gran mayoría de las personas sufrieron en sociedades e instituciones sociales injustas, insalubres, represivas y opresivas. Estas sociedades tradicionales eran censurables debido a su arraigada desigualdad, explotación y tradicionalismo irracional... Pero la noble aspiración humana a la autonomía, la igualdad y la prosperidad luchó poderosamente contra las fuerzas de la miseria y la opresión, y finalmente logró establecer sociedades modernas, liberales, democráticas, capitalistas y de bienestar. [Sin embargo] queda mucho trabajo por hacer para desmantelar los poderosos vestigios de desigualdad, explotación y represión. Esta lucha... es la única misión a la que realmente vale la pena dedicar la vida.

Claramente, esta narrativa se basa en gran medida en los fundamentos de Daño y Equidad para contar una historia de triunfo. Rechaza explícitamente el fundamento de Autoridad como una fuente de valor, retratando a la autoridad y su consiguiente valoración de la tradición como la causa raíz del mal en el mundo.

En contraste, la narrativa de la "comunidad perdida" es políticamente más conservadora. Se basa principalmente en los fundamentos del Intragrupo y la Autoridad para contar una historia de declive y decadencia:

> Érase una vez, cuando la gente vivía junta en comunidades locales, cara a cara, donde todos nos conocíamos y nos cuidábamos... la vida estaba firmemente entrelazada en

CC BY-NC-ND

> tejidos de cultura, fe y tradición, orgánicos e integrados. Realmente sabíamos quiénes éramos y sentíamos profundamente nuestra tierra, nuestros parientes y nuestras costumbres. Pero entonces sucedió algo terrible: la comunidad popular fue invadida por las barbaridades de la industria moderna, la urbanización, la racionalidad, la ciencia, la fragmentación, el anonimato… La fe comenzó a erosionarse, la confianza social se disipó, las costumbres populares se desvanecieron…. Todo lo que queda hoy son vestigios andrajosos de un mundo que hemos perdido. La tarea de quienes ahora lo ven claramente, es conmemorar y celebrar la comunidad popular, llorar su ruina y resistir, denunciar las depravaciones del racionalismo científico moderno que podría matar al Espíritu Humano.

Ninguna narrativa es correcta en ningún sentido objetivo. Ambas son formas en que los sociólogos han tratado de darle sentido a la historia. Como si emplearan dos cocinas diferentes, los académicos liberales, por un lado, y los conservadores, por otro, combinan y recombinan ingeniosamente sus elementos favoritos. Sin los cinco fundamentos dados de forma innata, no podría haber narrativas morales emocionalmente convincentes. Pero sin narrativa, nuestros conceptos morales estarían inconexos y serían difíciles de integrar en planes de acción coherentes.

Hemos incluido este extenso análisis de la narrativa en nuestro ensayo sobre el innatismo moral por dos razones principales. En primer lugar, la narrativa es una herramienta cultural importante para la modificación y socialización de las intuiciones fundamentales que están en el centro de este ensayo. La narración de historias es una parte indispensable de la educación moral en todas las culturas, e incluso el discurso moral de los adultos con frecuencia vuelve a apelar a narrativas como un medio para reclamar autoridad. En segundo lugar, como deja en claro la referencia a Bruner, el pensamiento narrativo en sí mismo es innato y un aspecto fundamental de nuestra arquitectura cognitiva. Es tan fundamental, al menos, como el modo de pensamiento 'proposicional'. De hecho, resulta plausible sostener que la moral y la capacidad humana para la narratividad han evolucionado conjuntamente, reforzándose mutuamente en nuestro reciente desarrollo filogenético.

7. Conclusión

de Waal (1996) sugiere que un elemento fundamental de la moral humana visible en los chimpancés es el deseo de paz y armonía dentro del grupo. Las celebraciones estallan cuando se resuelven las luchas de poder de larga data. Creemos que este deseo está relacionado con el fundamento Intragrupal: las criaturas que viven en grupo prefieren (tienen una tendencia innata a valorar) la armonía dentro de grupos cooperativos de los que dependen, tanto para su sustento material como para la defensa intergrupal. Encontramos este deseo en nosotros mismos: somos parte de la comunidad de investigadores de la moral que durante mucho tiempo ha estado dividida sobre la cuestión de su carácter innato. Esta división nos hace sentir incómodos, por el hecho de que realmente apreciamos y valoramos a los muchos miembros

CC BY-NC-ND

de nuestra comunidad, y, en este ensayo, hemos intentado mostrar cómo todos tienen razón en algo, cómo todos tienen algo que contribuir. Nuestra propuesta es que se podría restaurar cierto grado de armonía y sinergia, si la mayoría de los investigadores de la moral están dispuestos a respaldar esta afirmación: *el primer borrador de la mente moral tiene un contenido moral diverso que se especificó antes de la experiencia, pero este contenido dado de manera innata es revisado y ampliamente extendido durante el curso del desarrollo, a medida que los niños construyen activamente su conocimiento moral dentro de un contexto cultural que usa la narrativa para dar forma y guiar el desarrollo de virtudes específicas*, ¿alguien está listo para celebrar con nosotros o para proponer una declaración de consenso alternativa?

Agradecimientos

Agradecemos a Peter Carruthers, Jesse Graham, Angel Lillard, Shige Oishi y Dan Sperber por sus útiles comentarios sobre el primer borrador.

Referencias bibliográficas

Abu-Lughod, L. (1986). *Veiled sentiments*. Berkeley: University of California Press.

Aristotle. (1941). Nichomachean ethics. In R. McKeon (Ed.), *The basic works of Aristotle*, pp. 927-1112. New York: Random House.

Barkow, J. H., Cosmides, L., Tooby, J. (Eds.) (1992). *The Adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford.

Baron, Jonathan (1993). *Morality and rational choice.* Springer-Verlag.

Batson, C. D., O'Quinn, K., Fulty, J., Vanderplass, M., Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of Personality and Social Psychology, 45*, 706-718.

Boehm, C. (1999). *Hierarchy in the forest: The evolution of egalitarian behavior*. Cambridge, MA: Harvard University Press.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic.

Bruner, Jerome (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Buller, D. J. (2005). *Adapting minds: Evolutionary psychology and the persistent quest for human nature*. Cambridge, MA: MIT Press/Bradford Books.

Buller, D. J., Hardcastle, V. G. (2000). Evolutionary psychology, meet developmental neurobiology: Against promiscuous modularity. *Brain and Mind*, 1, 307-325.

CC BY-NC-ND

Burns, K. (1997). Lewis & Clark - The Journey of the Corps of Discovery. (Movie).

Buss, D. M. (2004). *Evolutionary psychology: The new science of the mind.* Boston: Allyn & Bacon.

Carruthers, Michael (1991). Narrativity: Mind-reading and making societies. In Andrew Whiten (Ed.), *Natural theories of mind: Evolution, Development and Simulation of Everyday Mindreading*, pp. 317-331. Oxford: Basil Blackwell.

Carruthers, P. (n/d). Moderately massive modularity. Available on the Internet at http://www.philosophy.umd.edu/Faculty/pcarruthers/Moderate-modularity.htm.

Casebeer, W. D. (2003). *Natural ethical facts: evolution, connectionism, and moral cognition*. Cambridge, MA: MIT Press.

Cervone, D., Shoda, Y. (Eds.) (1999). *The coherence of personality: Social-cognitive bases of consistency, variability, and organization*. New York: Guilford Press.

Churchland, P. M. (1996). *The engine of reason, the seat of the soul: a philosophical journey into the brain.* Cambridge, MA: MIT Press.

Churchland, P. M. (1998). Toward a Cognitive Neurobiology of the Moral Virtues. *Topoi*, 17, 83-96.

Clark, A. (2000). Word and action: reconciling rules and know-how in moral cognition. In R. Campbell and B. Hunter, eds., *Moral Epistemology Naturalized*, *Canadian Journal of Philosophy*, supplementary volume XXVI.

Coulter, A. (2003). *Treason: Liberal treachery from the cold war to the war on terror*. Crown.

de Waal, F. (1996). *Good natured: The origins of right and wrong in humans and other animals*. Cambridge, MA: Harvard University Press.

Darley, J. M., Batson, C. D. (1973). From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior. *Journal of Personality & Social Psychology, 27*, 100-108.

Dewey, John (1922). *Human Nature and Conduct: An Introduction to Social Psychology*. New York: Holt.

Doris, J. (1998). Persons, situations, and virtue ethics. *Nous, 32*(4), 504-530.

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral & Brain Sciences*, 16, 681-735.

Eliade, M. (1957/1959). *The sacred and the profane: The nature of religion*. San Diego: Harcourt Brace.

Fiske, A. P. (1991). *Structures of social life*. New York: Free Press.

CC BY-NC-ND

Fiske, A. P. (1992). Four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*, 689-723.

Fiske, A. P. (2004). Relational Models Theory 2.0. In N. Haslam (Ed.), *Relational models theory: A contemporary overview*, pp. 3-25. Mahwah, NJ: Lawrence Erlbaum.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Greene, J., Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences, 6*, 517-523.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., Cohen, J. D. (2001). An fMRI study of emotional engagement in moral judgment. *Science, 293*, 2105-2108.

Griffiths, P. (2002). What is innateness? *The Monist*, *85*(1), 70-85.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814-834.

Haidt, J., Bjorklund, F. (2008). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 2. The cognitive science of morality: Intuition and diversity*, pp. 181-217. Cambridge, Massachusetts: MIT Press.

Haidt, J., Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*, 98-116.

Haidt, J., Graham, J. (2009). Planet of the Durkheimians, where community, authority, and sacredness are foundations of morality. In J. T. Jost, A. C. Kay, & H. Thorisdottir (Eds.), *Social and psychological bases of ideology and system justification*, pp. 371-401. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195320916.003.015

Haidt, J., Hersh, M. A. (2001). Sexual morality: The cultures and reasons of liberals and conservatives. *Journal of Applied Social Psychology, 31*, 191-221.

Haidt, J., Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, Fall*, 55-66.

Haidt, J., Koller, S., Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613-628.

Haidt, J., Rozin, P., McCauley, C. R., Imada, S. (1997). Body, psyche, and culture: The relationship between disgust and morality. *Psychology and Developing Societies, 9*, 107-131.

Hampshire, S. (1982). Morality and convention. In A. Sen and B. Williams (Eds.), *Utilitarianism and Beyond*, pp. 145-158. Cambridge: Cambridge University Press.

CC BY-NC-ND

Hansen, C. (1991). Classical Chinese Ethics. In P. Singer (Ed.), *A companion to ethics*, pp. 69-81. Oxford, UK: Oxford.

Harman, G. (1999). Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, *99,* 315-32.

Hartshorne, H., May, M. A. (1928). *Studies in the nature of character. Vol. 1: Studies in deceit*. New York: MacMillan.

Haslam, N. (1997). Four grammars for primate social relations. In J. A. Simpson & D. T. Kenrick (Eds.), *Evolutionary social psychology*, pp. 297-316. Mahwah, NJ: Lawrence Erlbaum.

Hauser, M. (2006). *Moral Minds: How nature designed our universal sense of right and wrong*. New York: Harper Collins.

Jensen, L. A. (1997). Culture wars: American moral divisions across the adult lifespan. *Journal of Adult Development, 4*, 107-121.

Jensen, L. A. (1998). Moral divisions within countries between orthodoxy and progressivism: India and the United States. *Journal for the Scientific Study of Religion, 37*, 90-107.

Kelly, D., Stich, S. (2007). Two Theories About the Cognitive Architecture Underlying Morality. En P. Carruthers, S. Laurence y S. Stich, (Eds.), *The Innate Mind, Vol. 3 Foundations and the Future*. New York: Oxford University Press.

Kelly, D., Stich, S., Haley, K., Eng, S., Fessler, D. (2007). Harm, affect, and the moral/conventional distinction. *Mind and Language*, *22*(2), 117-131.

Kohlberg, L. (1969). Stage and Sequence: The Cognitive Developmental Approach to Socialization. In D. Goslin, (Ed.), *Handbook of Socialization Theory and Research*, pp. 347-480. Chicago, IL: Rand McNally.

MacIntyre, A. (1981). *After virtue*. Notre Dame: University of Notre Dame Press.

Marler, P. (1991). The instinct to learn. In Susan Carey and Rochel Gelman (eds.), *The epigenesis of mind*, pp. 37-66. Hillsdale: Erlbaum.

Marcus, G. (2004). *The birth of the mind*. New York: Basic Books.

May, L., Friedman, M., Clark, A. (Eds.). (1995). *Mind and morals: Essays on ethics and cognitive science*. Cambridge, MA: MIT Press.

McDowell, J. (1979). Virtue and reason. *Monist 62*, 331-350.

Meigs, A. (1984). *Food, sex, and pollution: A New Guinea religion*. New Brunswick, NJ: Rutgers University Press.

Merritt, M. (2000). Virtue ethics and situationist personality psychology. *Ethical theory and moral practice*, 3, 365-83.

CC BY-NC-ND

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Nussbaum, M. C. (1993). Non-relative virtues: an Aristotelian approach. In M. C. Nussbaum & A. Sen (Eds.), *The quality of life*, pp. 242-269. New York: Oxford University Press.

Redding, R. E. (2001). Sociopolitical diversity in psychology: The case for pluralism. *American Psychologist, 56*, 205-215.

Robert A. Pape (2005). *Dying to win: The strategic logic of suicide terrorism.* New York: Random House.

Parish, S. M. (1994). *Moral knowing in a Hindu sacred city*. New York: Columbia University Press.

Petrinovich, L., O'Neill, P., Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology, 64*, 467-478.

Piaget, J. (1965/1932). *The moral judgement of the child*. New York: Free Press.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Pincoffs, E. L. (1986). *Quandaries and virtues: Against reductivism in ethics*. Lawrence, Kansas: University of Kansas.

Richerson, P. J., Boyd, R. (1998). The evolution of human ultra-sociality. In I. Eibl-Eibesfeldt & F. K. Salter (Eds.), Indoctrinability, ideology, and warfare: Evolutionary perspectives, pp. 71-95. New York: Berghahn.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Ross, L., Nisbett, R. E. (1991). *The person and the situation*. New York: McGraw-Hill.

Rozin, P., Fallon, A. (1987). A perspective on disgust. *Psychological Review, 94*, 23-41.

Rozin, P., Haidt, J., McCauley, C. R. (2000). Disgust. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions,* pp. 637-653. New York: Guilford Press.

Samuels, R. (2004). Innateness in cognitive science. *Trends in Cognitive Sciences, 8*, 136-141.

Scanlon, T. (1998). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy, 2*, 307-320.

Shweder, R. (1990). In defense of moral realism: Reply to Gabennesch. *Child Development, 61*, 2060-2067.

Shweder, R. A. (1990). Cultural psychology: What is it? In J. W. Stigler, R. A. Shweder & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development*, pp. 1-43. New York: Cambridge University Press.

CC BY-NC-ND

Shweder, R. A., Balle-Jensen, L., Goldstein, W. M. (1995). Who sleeps by whom revisited: A method for extracting the moral goods implicit in practices. In J. Goodnow, P. Miller & F. Kessel (Eds.), *Cultural practices as contexts for development*, pp. 21-39. San Francisco: Jossey-Bass.

Shweder, R. A., Mahapatra, M., Miller, J. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children*, pp. 1-83. Chicago: University of Chicago Press.

Shweder, R. A., Much, N. C. (1991). Determinations of meaning: Discourse and moral socialization. In R. A. Shweder (Ed.), *Thinking through cultures*, pp. 186-240. Cambridge, MA: Harvard University Press.

Shweder, R. A., Much, N. C., Mahapatra, M., Park, L. (1997). The "big three" of morality (autonomy, community, and divinity), and the "big three" explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health*, pp. 119-169. New York: Routledge.

Smith, C. (2003). *Moral, believing animals: Human personhood and culture*. Oxford, UK: Oxford.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*, pp. 39-67. Cambridge, UK: Cambridge University Press.

Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Structure and contents*, pp. 53-68. New York: Oxford.

Sperber, D., Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences, 8*, 40-46.

Sreenivasan, G, (2002). Errors about errors: Virtue theory and trait attribution. *Mind, 111*, 47-68.

Sripada, C. S., Stich, S. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind Vol. 2. Culture and cognition*, pp. 280-301. Oxford: Oxford University Press.

Tajfel, H., Billig, M. G., Bundy, R. P., Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149-178.

Tooby, J., Cosmides, L., Barrett, C. (2005). Resolving the debate on innate ideas: Learnability constraints and the evolved interpenetration of motivational and conceptual functions. In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Structure and contents*, pp. 305-337. New York: Oxford.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35-57.

CC BY-NC-ND

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, England: Cambridge University Press.

Turiel, E., Killen, M., Helwig, C. C. (1987). Morality: Its structure, function, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children*, pp. 155-243. Chicago: University of Chicago Press.

Valdesolo, P., DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science, 17*, 476-477.

Wimsatt, W. C. (1999). Generativity, entrenchment, evolution, and innateness: Philosophy, evolutionary biology, and the conceptual foundations of science. In V. G. Hardcastle (Ed.), *Where biology meets psychology: Philosophical essays*, pp. 139-180. Cambridge, MA: MIT Press.

Zahn-Waxler, C., Radke-Yarrow, M., King, R. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child Development, 50*, 319-330.

CC BY-NC-ND