

La noción de daño y su rol en los juicios morales. Un debate abierto^{*,**}

The Notion of Harm and its Role in Moral Judgments. An Open Debate

GUSTAVO SILVA^{***}
GUSTAVO REYES^{****}
GUSTAVO PEÑA^{*****}
ALEJANDRO ROSAS^{*****}

Resumen: Investigadores pertenecientes a la tradición de la psicología del desarrollo cognitivo han llevado a cabo estudios que sugieren que existe un vínculo entre la percepción de daño y el dominio moral. Frente a esta propuesta unificadora del dominio moral han surgido críticas desde la psicología cultural. Haidt publicó en 1993 uno de los estudios más influyentes en esta línea, en el que afirmaba haber encontrado evidencia que sugiere la existencia de dominios morales no vinculados a la noción de daño. En este artículo se hará una presentación general del debate en torno al contenido del dominio moral. Luego plantearemos y evaluaremos el estudio de Haidt et al. (1993). Por último, expondremos resultados experimentales que cuestionan las conclusiones de ese estudio y sugieren que el debate sigue abierto.

Palabras clave: dominio moral, normas morales, daño, psicología cognitiva.

Abstract: Researchers belonging to the tradition of cognitive-developmental psychology have suggested in their studies that the perception of harm is central to the moral domain. This unifying proposal has provoked the criticisms of researchers working in Cultural Psychology. Haidt published in 1993 a very influential study, claiming to have found evidence for the existence of moral domains not linked to the notion of harm. In this paper, we briefly present the debate around the content of the moral domains. We then critically examine the study by Haidt et al. (1993). Finally, we present experimental results that question the conclusions of that study and suggest that the debate is still open.

Keywords: moral domain, harm, moral norms, cognitive psychology.

Recibido: 13/12/2018. Aceptado: 13/06/2019.

* Esta investigación contó con el apoyo financiero de la Vicerrectoría de Investigación de la Universidad Nacional de Colombia, mediante el proyecto 37159, 2017-2018.

** Los autores reconocen y agradecen la cuidadosa labor de los revisores anónimos designados por la Revista.

*** Profesor, Departamento de Humanidades, Universidad El Bosque y candidato a doctor en filosofía, Universidad Nacional de Colombia. Email: gasilvac@unal.edu.co

**** Filósofo, egresado del programa de Filosofía, Universidad Nacional de Colombia.
Email: gareyesh@unal.edu.co

***** Candidato a Magíster en Filosofía, Universidad Nacional de Colombia. Email: gpenac@unal.edu.co

***** PhD, Profesor Titular, Departamento de Filosofía, Universidad Nacional de Colombia.
Email: arosasl@unal.edu.co

1. Introducción

Cuando pensamos en las normas que regulan el comportamiento humano, es común considerar dos tipos de normas: normas morales y normas sociales convencionales. Solemos juzgar que las normas que prohíben el asesinato, el robo o el engaño son normas morales; mientras que consideramos normas sobre vestimenta o referidas a comportamientos de etiqueta como asuntos de convención social. A pesar de que realizamos muchos juicios de este tipo, no hay claridad respecto a qué características distinguen una norma moral de otras normas como la de etiqueta. ¿Bajo qué criterios hacemos esta distinción? Este problema ha sido el origen de un debate interdisciplinar con la participación de la psicología, la filosofía, la antropología, las neurociencias, entre otras. El debate busca determinar si es posible o no identificar una característica que distinga las normas morales de otros tipos de normas, de modo que el dominio moral en general quede unificado bajo un único contenido (Rosas, 2013, 494).

El estudio de Haidt, Koller y Dias (1993) pareció ofrecer la evidencia definitiva a favor de la tesis del contenido irreductiblemente plural del dominio moral. Ese trabajo pretende haber demostrado que hay “males sin daño”. Sin embargo, creemos que sus análisis no permiten aún cerrar el debate, como intentaremos mostrar a continuación.

En el presente artículo evaluaremos los análisis presentados por Haidt et al. (1993) a favor de la teoría de dominios morales distintos e independientes y, en especial, a favor de la existencia de acciones clasificadas como inmorales a pesar de no causar daño a otros. Con el fin de dar apoyo experimental a nuestra evaluación crítica, presentamos los resultados de una réplica parcial del trabajo de Haidt et al. (1993), utilizando las dos viñetas de violaciones de pureza que sobresalieron en su estudio y que describen conductas relativas a la sexualidad. Nuestros resultados muestran correlaciones positivas significativas entre la atribución de daño a otros por esas acciones impuras y la moralización de las mismas.

Aunque para el psicólogo cultural y sus colaboradores las situaciones diseñadas en el laboratorio sean “males sin daño”, nuestros resultados indican que son percibidas como dañinas por los individuos que las condenan. Los resultados sugieren que la hipótesis de Haidt y colaboradores no puede considerarse comprobada experimentalmente por ese estudio.

2. La hipótesis del daño

En los años 70 del siglo pasado, Elliot Turiel desarrolló un experimento conocido como la prueba moral-convencional (*moral-conventional task*). El experimento es importante dentro del debate sobre la posible unificación del dominio moral bajo un único contenido. Consiste en presentar a los sujetos experimentales un conjunto de historias que describen transgresiones de normas consideradas comúnmente como morales o como convencionales. A cada una de estas historias le sigue una serie de preguntas encaminadas a determinar: a) cómo juzgan los sujetos estas acciones –si las consideran incorrectas y, de ser así, qué tanto–; b) qué características atribuyen a una norma que proscribe tal acción –si juzgan tal norma como dependiente o no de una autoridad y si consideran que la norma es válida o no para todo tiempo y lugar (universalidad)– y c) cómo justifican dichas normas. Los

datos obtenidos mediante estas preguntas permiten saber cómo los sujetos caracterizan cada caso y si tal caracterización muestra una distinción entre el dominio moral y el convencional. Con los resultados de estos experimentos Turiel (1983, 1998) desarrolló su propuesta sobre el daño, la justicia y el derecho como contenidos nucleares y unificadores del dominio moral.

Siguiendo a Daniel Kelly y colegas (2007), las conclusiones de Turiel se pueden resumir en tres puntos:¹

C1. En los experimentos (que aplican la prueba moral-convencional) los sujetos típicamente exhiben uno de dos patrones de respuesta. En el patrón de respuesta moral, las normas son juzgadas como independientes de autoridades y generales en alcance (i.e., universales); las violaciones (de estas normas) son más graves y las normas son justificadas apelando a las nociones de daño, justicia o derechos. En el patrón de respuesta convencional las normas son juzgadas como dependientes de autoridades y con un alcance local (i.e., no universales); las violaciones de dichas normas son menos graves y, adicionalmente, las normas no son justificadas apelando a las nociones de daño, justicia o derechos.

C2. a) Transgresiones que involucran daño, justicia o derechos evocan un patrón de respuesta moral. b) Transgresiones que no involucran daño, justicia o derechos evocan un patrón de respuesta convencional.

C3. Las regularidades descritas en C1 y en C2 son panculturales y emergen a edad temprana (4 años) en el desarrollo.

La descripción expuesta en C1 hace referencia a características o propiedades de tipo *formal*. Así, desde la tradición del desarrollo cognitivo se postula que el dominio moral está formalmente unificado por las características de universalidad e independencia de autoridades. Adicionalmente, se considera que las acciones que violan dichas normas merecen castigos severos y acciones disuasivas. En lo que sigue, nos referiremos a la atribución de estas propiedades formales como *moralización*.

Desde la tradición de la psicología cognitiva se ha propuesto que las nociones de daño, justicia y derechos son centrales y unifican el contenido de la moralidad. En este sentido, C1-C3 nos permiten especificar esta propuesta y plantearla del siguiente modo: las nociones de daño, justicia y derechos unifican al dominio moral tanto de manera causal como conceptual. El vínculo causal es sugerido por C2, en donde se indica que *solamente* los casos que involucran daño (consecuencias dañinas), injusticia y violación de derechos causan los patrones de moralización. El vínculo conceptual es sugerido en C1: *solamente* son morales aquellas normas que proscriben acciones que causan daño, cometen injusticia y violan derechos. Esto sugiere que las personas piensan en el concepto de *norma moral* como asociado a estas nociones. C3 indica que todos estos vínculos son panculturales. De ahí que se considere que esta descripción abarca el dominio moral en general.

Turiel (1983) describe el contenido del dominio moral en términos de daño, justicia y derechos. Sin embargo, describir el dominio moral mediante los conceptos de justicia y derechos puede implicar circularidad. Considérese, por ejemplo, la noción de derechos inalienables. Por definición, un derecho inalienable es universal y válido con independencia del juicio de

1 Esta es una ligera adaptación del sumario que D. Kelly y sus colegas presentan en el artículo *Harm, Affect, and the Moral/Conventional Distinction* (Kelly et al., 2007, 119-120).

una autoridad. Describir el contenido del dominio moral mediante esta noción convierte la conexión entre lo moral y las propiedades formales indicadas en C1 (*moralización*) en una tautología que no requiere confirmación empírica (Stich et al., 2009, 95-96), pues los derechos inalienables son universales e independientes de autoridades. Una descripción del contenido del dominio moral basada solo en la noción de daño no presenta este problema. El vínculo entre propiedades morales formales y la noción de daño parece solo ser verificable de modo empírico. Con miras a investigar estas relaciones, resulta más conveniente describir el contenido del dominio moral mediante la noción de daño.

Considerando lo anterior, formularemos la propuesta de la psicología del desarrollo cognitivo del siguiente modo: el dominio moral puede ser descrito en dos aspectos, uno formal y otro de contenido. Formalmente, el dominio moral está descrito y unificado por las propiedades de universalidad e independencia de autoridades. Respecto a su contenido, el dominio moral está unificado por la noción de daño. Esta noción se vincula al dominio moral de manera causal y conceptual. A esta hipótesis unificadora del contenido moral nos referiremos como *hipótesis del daño*. El término “daño” puede referirse a muchas cosas, pero aquí usaremos la siguiente definición: “daño” se refiere al empeoramiento intencional del bienestar físico/material o psicológico de una persona por acción de otra persona.

Tanto la hipótesis del daño como la descripción formal ya expuesta han sido objeto de diversas réplicas. El debate ha sido motivado principalmente por la afirmación de pan-culturalidad (C3). Prácticamente no se pone en duda que los planteamientos de C1 y C2 sean ciertos para dominios morales particulares de ciertos grupos o culturas, principalmente de Occidente. Sin embargo, muchos investigadores, en su mayoría pertenecientes a la tradición de la psicología cultural, afirman que esta descripción no abarca el dominio moral en general.

A propósito de este problema se ha fortalecido en los últimos años una propuesta de corte culturalista. Esta plantea que, aunque se acepte que la percepción de daño es importante para reconocer las acciones morales, la conformación de los juicios morales depende de más contenidos básicos, independientes del daño a otras personas (Cf. Flanagan, 2017).

Las principales críticas a la hipótesis del daño están basadas en experimentos que usan los criterios morales de la prueba moral-convencional para identificar las normas morales (Rosas, 2013, 494). Uno de los primeros estudios en contra de la hipótesis de daño fue llevado a cabo por Shweder, Mahapatra y Miller (1987). Ellos reportaron que sujetos en India moralizan normas sobre comportamientos referidos a asuntos de castidad, jerarquía social y pureza que no fueron moralizados por sujetos en Estados Unidos porque, según los autores, en India los sujetos moralizan conductas que, aparentemente, no involucran daño. Y lo mismo parece ocurrir en otras culturas (Miller, Bersoff y Harwood, 1990; Shweder, 1990; Shweder y Sullivan, 1993).

En este sentido, Shweder ha propuesto una clásica división del dominio moral en *ética de la autonomía*, *ética de la comunidad* y *ética de la divinidad*. El primer caso se refiere a códigos cuyo objeto es proteger e incrementar la autonomía del individuo, sus elecciones libres y el autocontrol (Shweder, 1990, citado por Haidt et al., 1993). Este tipo de códigos fortalecen la autonomía y pueden relacionarse con normas que prohíben el daño a otros. Sin embargo, en la visión transcultural se pueden reconocer otros dos códigos morales (Shweder, 1990). Uno de ellos es la *ética de la comunidad* que conceptualiza al individuo

como actor de una empresa colectiva, donde debe desempeñar un papel en interdependencia con los demás. Así, el código de la comunidad se guía por el respeto, el deber, la obediencia a la autoridad y las acciones consistentes con la función y la posición social del individuo. La *ética de la divinidad*, a su vez, asume a la persona como una entidad espiritual obligada a cuidar la pureza y la santidad del espíritu (Haidt et al., 1993, 614). Para Shweder (1990), el dominio de la moralidad ha sido equivocadamente restringido a lo que él denomina *ética de la autonomía*, sobre todo porque los estudios empíricos de esta teoría se han concentrado en poblaciones occidentales (regularmente estadounidenses) y han olvidado la diversidad cultural de otras regiones del mundo. Por ejemplo, Shweder et al. (1987), resaltan que los discursos morales en India incluyen consideraciones muy distintas a las que son típicas en la cultura occidental.

Las afirmaciones de Shweder et al. (1987) fueron cuestionadas por los psicólogos del desarrollo cognitivo. Turiel y otros han argumentado que las diferencias entre norteamericanos e indios encontradas por Shweder no se deben a diferencias sustanciales en el contenido del dominio moral, sino a diferencias en las creencias factuales sobre las consecuencias de actos contra la castidad, la pureza, las jerarquías endogrupales y la lealtad. El estudio de Shweder et al. (1987) no investiga la posibilidad de que los sujetos de India moralicen las violaciones de las normas sobre estos actos porque consideran que las transgresiones a estas normas tienen consecuencias dañinas, i.e., involucran daño a otros individuos o incluso a grupos de ellos.

En medio de esta controversia teórica que presenta a Turiel y Shweder como los representantes en orillas opuestas del debate, aparece Jonathan Haidt, filósofo y psicólogo de la Universidad de Nueva York, que intenta desde mediados de la década de los 90 del siglo pasado cerrar dicha controversia a favor de una teoría enmarcada en el culturalismo de Shweder. Haidt se apoya en los resultados experimentales obtenidos con sus colaboradores en las tres últimas décadas (Haidt et al., 1993; 2001; 2007). Su propuesta postula cinco categorías o *fundamentos morales* que dan lugar a otras tantas categorías de juicios intuitivos. Estos fundamentos de la moral serían universales, dado que son producto de nuestra evolución como especie. Desde esta perspectiva los juicios morales pueden generarse sin tener en cuenta contenidos referidos al daño. Haidt se apoya en resultados experimentales que sugieren la existencia de conductas que son moralizadas aunque no produzcan daño, especialmente hacia otros. Estos ejemplos de moralización son un contraejemplo a los hallazgos de Turiel, pues se moralizan conductas *sin efectos dañinos*, demostrando con esto que el daño no es el único contenido básico de la moral.

Haidt et al. (1993) pretenden haber mostrado experimentalmente que personas de estratos socioeconómicos bajos de ciudades en Brasil y Estados Unidos moralizan actos impuros, aun sin verlos dañinos. Haidt y sus colaboradoras denominan a estos actos “males sin daño”. Las pruebas experimentales parecen contundentes, pero en nuestra opinión hizo falta presentar un análisis de correlación entre la percepción de daño de un acto impuro *X* y la moralización de ese mismo acto *X*. Solo así podemos saber si los participantes, mayoritariamente, asocian su juicio moral sobre *X* con la percepción que tienen del daño producido por *X*. Si hay una correlación fuerte, quiere decir que los participantes asocian su atribución de daño y su moralización. El resultado contrario es el que se requeriría para dar apoyo a la Teoría de los Fundamentos Morales de Haidt.

3. La propuesta de Haidt. El irrespeto, la molestia y el daño como factores moralizadores independientes

Es importante señalar que Haidt no niega que el daño sea parte del dominio moral (Haidt et al., 1993, 613). Lo que cuestionan sus conclusiones es que el dominio moral en general se agote o esté unificado bajo la noción de daño.

El estudio de 1993 que examinamos aquí críticamente se realizó con 180 adultos y 180 niños de bajo y alto nivel socioeconómico (B-SES y A-SES, respectivamente), pertenecientes a tres ciudades: Porto Alegre y Recife, ambas de Brasil, y Filadelfia en EEUU². La clasificación por ciudad y estatus socioeconómico permitió agrupar a los sujetos en seis *grupos culturales* (estatus socioeconómico: 2 × ciudades: 3). Este experimento fue novedoso dentro del debate de descripción y unificación del dominio moral, porque se centró en investigar la moralización de transgresiones ofensivas, aunque carentes de consecuencias dañinas directas —“males sin daño”—. Las transgresiones descritas son, en principio, carentes de consecuencias dañinas porque son desarrolladas en privado. Sin embargo, Haidt y sus colaboradoras reconocen que si estas transgresiones fueran realizadas en público, causarían daño a sus observadores (Haidt et al., 1993, 615).

Las transgresiones elegidas para este estudio incluyen actos comúnmente considerados como molestos por ser irrespetuosos o repugnantes. Los estímulos consisten en cinco historias breves, a cada una de las cuales sigue una batería de preguntas. Dos de las historias describen comportamientos irrespetuosos desarrollados en privado: una mujer que limpia su baño con una vieja bandera de su país (Brasil o EE. UU.) y un hombre que, por estar ocupado, no cumple una promesa que hizo a su madre en su lecho de muerte: que le visitaría en su tumba cada semana. Las otras tres historias describen actos privados comúnmente considerados como repugnantes: una familia que cocina y se come a su perro luego de que este muere atropellado accidentalmente; una pareja de hermanos que, cuando nadie está cerca, se besan apasionadamente en la boca y un hombre que semanalmente compra un pollo muerto en un supermercado para cocinarlo y comérselo solo, pero que primero lo usa como juguete sexual³.

Las preguntas planteadas después de cada una de estas historias, son las siguientes (Haidt et al., 1993, 617):

- a. *Evaluación*: ¿Qué piensas acerca de esto, está muy mal, un poco mal o perfectamente bien realizar este [acto especificado]?
- b. *Justificación*: ¿Puedes decirme por qué?
- c. *Daño*: ¿Sufriría alguien un daño a causa de lo que hace [el agente]? ¿Quién? ¿Cómo?

2 Dado que este estudio busca determinar si el dominio moral varía sustancialmente entre culturas, dos variables culturales fueron consideradas: estatus socioeconómico (SES) y nivel de *occidentalización* —estilo de vida basado en tradiciones europeas—, lo cual incluye la estructura política democrática y la economía industrializada (Haidt et al., 1993, 615). Filadelfia y Recife mostraron las mayores diferencias en su nivel de occidentalización, siendo Recife la ciudad menos *occidentalizada* de las tres.

3 La historia del hombre que se masturba con un pollo muerto no fue presentada a los niños. Por lo tanto, los niños respondieron solo a cuatro historias. Para uno de los grupos de niños, la historia de los hermanos que se besan fue reemplazada por otra historia.

- d. *Molestia*: Imagina que tú realmente ves a alguien desarrollando este acto. ¿Te molestaría esto o no te importaría?
- e. *Interferencia*: ¿Debería el agente ser detenido o castigado de algún modo?
- f. *Universalidad*: Imagina dos países. En el país A, las personas hacen [acto específico] muy seguido, y en el país B, las personas nunca hacen ese [acto específico]. ¿Ambas costumbres están bien, o una de ellas está mal o es indebida?

Al considerar la crítica que Turiel planteó a Shweder, los investigadores (Haidt et al., 1993) incluyeron la pregunta *Daño*. La pregunta *Molestia* fue incluida para investigar el papel que las reacciones afectivas tienen en la moralización. Las preguntas *Interferencia* y *Universalidad* miden la moralización (por oposición a asuntos de fuero personal) que hacen o no los participantes de los actos estudiados. Una respuesta afirmativa a la pregunta *Interferencia* sugiere que el acto es considerado como transgresión de una norma moral. En la pregunta *Universalidad*, responder que una de las costumbres está mal o es indebida es un indicador de moralización.

Los investigadores eligieron historias de irrespeto o repugnancia, porque es común que comportamientos de este tipo sean juzgados en diversas culturas como transgresiones morales. No obstante, ellos mismos reconocen que es una cuestión empírica el determinar si estos comportamientos de irrespeto o repugnancia son juzgados como transgresiones morales por su potencial de daño o si son juzgados así por alguna otra razón independiente de las posibles consecuencias (dañinas) de estos actos (Haidt et al., 1993, 615).

Para probar su hipótesis, Haidt et al. (1993) necesitan encontrar evidencia de poblaciones con actitudes de moralización no vinculadas a la percepción de un daño. En las siguientes secciones, expondremos los resultados generales del experimento y nuestras reservas a varios de los análisis que los investigadores hacen para defender la existencia de “males morales sin daño”.

3.1. Resultados generales de Haidt et al. (1993)

El principal hallazgo de los investigadores fue una clara distinción entre los grupos de Alto Estrato Socioeconómico (A-ESE) y los de Bajo Estrato Socioeconómico (B-ESE). Los B-ESE resultaron menos permisivos –más moralizadores– que los A-ESE. En cuanto a la percepción de daño (pregunta *Daño*), todos los grupos mostraron un comportamiento similar, pues ninguna de las variables demográficas –ciudad, ESE y edad– mostró tener impacto (Haidt et al., 1993, 618, 620). Es de notar que la historia del beso apasionado entre hermanos y la del pollo como juguete sexual suscitaron los más altos porcentajes de percepción de daño entre adultos: 59% del grupo B-ESE y 45% del grupo A-ESE mencionaron algún daño como consecuencia de esas conductas, así como también los más altos porcentajes de molestia. Los adultos del grupo B-ESE mostraron mayores niveles de molestia que los adultos del grupo A-ESE (Haidt et al., 1993, 618). Los grupos culturales de niños no mostraron diferencias en cuanto a la pregunta *Molestia*; la mayoría de niños respondió afirmativamente a esta pregunta (Haidt et al., 1993, 620). Tras promediar resultados, los investigadores concluyeron que, *en general*, las historias fueron percibidas como carentes de daño y como ofensivas.

Ahora bien, respecto a actitudes moralizadoras, medidas por las preguntas *Interferencia* y *Universalidad*, surgió una clara distinción al considerar el Estrato Socioeconómico (ESE), mientras que la otra variable cultural (ciudad, que expresa el nivel de occidentalización), no mostró gran impacto en estas pruebas. Los porcentajes que los investigadores presentan (ver en su artículo las tablas 1 y 2) muestran que los grupos B-ESE moralizaron las cinco historias mucho más que los grupos A-ESE⁴. Basados en los resultados hasta ahora mencionados, los investigadores indicaron que los participantes de B-ESE moralizan a pesar de no percibir daño.

Ahora, considerando el tipo de datos presentados, la estrategia de Haidt et al. (1993) parece ser la siguiente: 1) Afirmar que las historias fueron *en general* percibidas como no dañinas y como molestas. 2) Mostrar que B-ESE moralizaron las historias, mientras que A-ESE no. 3) Sugerir a partir de 1) y 2) que B-ESE moralizan por algo diferente a un vínculo con daño, al parecer, por reacciones afectivas (molestia). Pero estas reflexiones no están basadas en un análisis de correlación entre las respuestas de un mismo sujeto a las preguntas *Daño*, *Molestia*, *Interferencia* y *Universalidad*, separadamente para cada una de las historias. Desde nuestro punto de vista, los investigadores basan su argumento en generalizaciones que corren el riesgo de invisibilizar relaciones importantes. Habría sido más apropiado, incluso indispensable, a nuestro entender, presentar los análisis de correlación y regresión entre la variable independiente que representa la percepción individual de daño para una acción específica y la variable dependiente “moralización”.

3.2. *El análisis de concordancia*

Haidt et al. (1993, 624) presentan sus resultados más pertinentes como un análisis que evalúa la concordancia, para cada sujeto, entre datos de *Daño-Moralización* y *Molestia-Moralización*. Mediante este análisis, los investigadores quisieron determinar qué variable predice mejor la moralización: el daño o la molestia. Las “concordancias” fueron evaluadas del siguiente modo:

- *Daño-Moralización*: se midió en cuántas de las cinco historias se percibió daño (medido por pregunta *Daño*) y se moralizó y en cuántas no hubo ni percepción de daño ni moralización.
- *Molestia-Moralización*: igual que arriba, se midió la concordancia entre sentir molestia/no sentirla y moralizar/no moralizar la acción impura.

Es decir, para cada uno de los noventa adultos de A-ESE se calculó el porcentaje de las cinco historias en las que el participante citó una víctima y moralizó o no citó víctima y no moralizó; luego se sacó un promedio de esos porcentajes para los noventa adultos. Pero en este análisis se invisibiliza por completo la frecuencia con la que los individuos de una muestra dada perciben daño y moralizan en cada historia en particular. Lo importante no es ver el promedio del porcentaje de concordancia, entendido como en cuántas de las

4 Un análisis que combina los datos obtenidos con ambas preguntas es presentado en las tablas de la página 623 de Haidt et al. (1993). Este análisis muestra el mismo resultado entre B-ESE y A-ESE.

cinco historias el participante atribuye daño y moraliza⁵ o no atribuye daño y no moraliza. Lo importante es ver, más bien, qué porcentaje del total de individuos ve daño y moraliza o no ve daño y no moraliza en cada una de las historias. Esto último tiene más sentido, pues las historias son distintas y los sujetos están contestando a las preguntas de daño y moralización historia por historia.

Con base en ese cálculo de las concordancias, los investigadores descubren que en el A-ESE las medias de porcentajes de concordancia entre daño y moralización fueron mayores que las medias entre molestia y moralización. En cambio, en el B-ESE las medias de porcentajes de concordancia fueron mayores para molestia que para moralización. Los investigadores concluyeron que en los grupos más severos moralmente (B-ESE) las actitudes de moralización no están estrechamente vinculadas a la noción de daño y pueden ser mejor predichas por la variable “molestia” que por la variable “daño”, es decir, preguntando “¿te molestaría ver tal acto?” más bien que “¿es alguien dañado por tal acto?” (Haidt et al., 1993, 624-625). En la segunda pregunta se presupone que el acto es privado, en la primera necesariamente se tiene que levantar esa restricción.

4. Consideraciones críticas: daño y molestia

Aun si, por hipótesis, aceptásemos el análisis de concordancia de Haidt et al. (1993), queda la duda del verdadero significado de la variable “molestia”. Haidt afirma que actos irrespetuosos pueden causar daño psicológico (emociones negativas) a quien los *presencia* (Haidt et al., 1993, 615). Sobre actos repugnantes, los investigadores también reconocen que estos pueden tener potencial de daño. Los actos que Haidt et al. (1993) usan como estímulos para el experimento son desarrollados en privado: “Las ofensas sin daño en la presente investigación son todas privadas y consensuadas” (Haidt et al 1993, 613). Por esa razón, la pregunta *Daño* no serviría, en principio, como un indicador de daño psicológico a espectadores, pues los actos privados los excluyen. En cambio, la pregunta *Molestia* rompe con la característica de privacidad y da, en principio, vía libre a los daños que esas conductas pueden causar en sus espectadores (daños que Haidt y colaboradores reconocen: Haidt et al., 1993, 615). La pregunta se plantea así: “Imagina que tú realmente ves a alguien realizando este acto. ¿Te molestaría esto o no te importaría?” (Haidt et al., 1993, 617). Nada en el experimento impide una interpretación de esta pregunta como un medidor de daño psicológico a espectadores. En consecuencia, en defensa de una moralidad basada en el daño, se podría decir que los grupos moralizadores –B-ESE–, en los que la moralización correlaciona más con los datos de *Molestia* que con los de *Daño*, moralizan por percibir daño psicológico a espectadores. Dada esta lectura, no parece muy plausible decir que la variable molestia mide algo diferente al daño. Faltaría por entender la razón de las diferencias encontradas por los investigadores entre grupos A-ESE y grupos B-ESE. Pero el objetivo central no parece haberse logrado: encontrar un ámbito de conductas que se moralizan a pesar de no percibirse en ellas efectos dañinos.

5 Haidt y colegas tampoco explican cómo deciden si alguien moraliza, pues teniendo dos preguntas con respuesta binaria sobre moralización, tendrían que explicar cómo computan el valor de “moralización” a partir de dos valores binarios iniciales que pueden ser contrarios.

5. Estudio correlacional: la percepción de daño como predictor de la moralización de las violaciones de pureza

Con el fin de re-examinar las relaciones entre percepción de daño y la moralización de los actos impuros, realizamos una réplica parcial del experimento de Haidt et al. (1993). Tuvi- mos en cuenta solo las viñetas en donde los agentes violan normas de pureza en conductas relacionadas con la sexualidad. La ausencia de un análisis de correlación entre percepción de daño y moralización fue la principal carencia encontrada en el estudio de Haidt et al. (1993). Esto fue, precisamente, lo que tratamos de subsanar con nuestro estudio.

5.1. Metodología

5.1.1. Participantes

Las encuestas se aplicaron de manera presencial, aunque sin solicitar ni registrar datos de identificación personal, a 185 estudiantes y trabajadores en la Universidad Nacional de Colombia. El 35% de los encuestados fueron mujeres (mujeres = 0; hombres = 1). Se preguntó por el Estrato Socioeconómico (ESE) con niveles que van del 1 (más bajo) al 6 (más alto). El promedio de edad fue de 24 años con una Desviación Estándar de 9.5.

5.1.2. Materiales y procedimientos

Se usaron tres situaciones como estímulos. Dos de ellas fueron calificadas en Haidt et al. (1993) como actos repugnantes en privado y, por tanto, sin daño a otros (el caso del pollo como juguete sexual y el caso del beso erótico entre hermanos). Utilizamos un tercer estímulo de control, en el que no se viola ninguna norma de pureza o de daño.⁶

Después de cada estímulo preguntamos a los participantes:

1. ¿Sufre alguien (persona distinta a [nombre del agente en el estímulo]), la comunidad o la sociedad algún daño por la acción privada de [nombre del agente en el estímulo]? A esta pregunta contestaron en una escala de 7 puntos, donde 1= ningún daño y 7 = muchísimo daño. (variable “atribución de daño”).

Luego preguntamos por la evaluación moral del acto, replicando las dos medidas de moralización utilizadas por Haidt et al. (1993):

2. ¿Crees que habría que impedir y/o castigar esa acción privada? SI___ NO___

La respuesta “Sí” indica moralización y fue codificada como 1, la respuesta “No”, como 0. (variable moralización1).

3. Supón que te enteras que en el país X la gente acostumbra hacer (a escondidas) actos de ese tipo mientras que en el país Y nunca acostumbran hacerlos. ¿Qué pensarías?

Una de ellas es mala o incorrecta___ Ambas costumbres son aceptables___

6 “Patricia, nuestra vecina, se viste a veces completamente de azul, y se mira en el espejo así vestida durante largo rato, sin que nadie la vea.”

La primera respuesta indica moralización y fue codificada como 1, la segunda respuesta como 0. (variable moralización2).

Computamos el promedio de atribución de daño a los dos estímulos de pureza como variable “Atribución Daño-pureza”, con 14 niveles. Asimismo, computamos el promedio de las dos preguntas de moralización, y luego el promedio de moralización para los dos estímulos de pureza en una sola variable “Moralización Pureza”, que resultó con 4 niveles: 0, 0.25, 0.5 y 1 (partiendo de dos valores 0 y 1). Hicimos el mismo cómputo para la moralización del único escenario control que resultó teniendo, lógicamente, tres niveles: 0, 0.5 y 1 (partiendo de dos valores 0 y 1).

Siendo nuestro objetivo establecer si la atribución de daño a los estímulos de pureza predice su moralización, analizamos la relación entre esas variables con correlaciones Pearson y regresiones lineales.

5.1.3. Resultados

A continuación presentamos la tabla con los estadísticos descriptivos para las variables demográficas y experimentales.

Tabla 1. Estadísticos descriptivos

	N	Mín.	Máx.	Media	Desviación Es.
Edad	167	16	61	24.32	9.507
Sexo	183	0	1	.65	.478
Estrato	183	1	6	2.71	1.010
Atrib Daño_control	184	1	7	1.14	.634
Moraliz_control	182	.0	1.0	.058	.1916
Atrib Daño_pureza	185	1	7	2.711	1.8360
Moraliz_pureza	185	.00	1.00	.4036	.37903
Valid N (listwise)	162				

Puede observarse que en el estímulo control, la media de la atribución de daño así como la media de la moralización están muy cercanas al mínimo, como es de esperarse de una acción neutra que no viola ninguna norma moral. En los estímulos que representan una acción impura, la media de atribución de daño está por debajo del punto medio en la escala de 7. La moralización de la acción impura está cercana al punto medio.

A continuación presentamos las correlaciones entre la variable dependiente (moralización) y las variables independientes (atribución de daño, edad, estrato y sexo).

Tabla 2. Correlaciones

		Moraliz Pureza	Moraliz Control	AtribDaño Control	AtribDaño Pureza	Edad	Estrato	Sexo
Moraliz Pureza	Pearson	1	.193**	.027	.524**	.419**	-.268**	-.239**
	Correlation							
	Sig. (2-tailed)		.009	.720	.000	.000	.000	.001
	N	185	182	184	185	167	183	183
Moraliz control	Pearson	.193**	1	.070	.152*	.179*	-.072	-.046
	Correlation							
	Sig. (2-tailed)	.009		.345	.041	.022	.336	.537
	N	182	182	182	182	164	180	180

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

En el caso de los estímulos de pureza se observa que la edad, el sexo, el estrato y la atribución de daño son todos predictores significativos. A mayor edad, más moralización; las mujeres moralizan más que los varones; mientras más alto es el ESE, menor moralización; a mayor atribución de daño, mayor moralización. El predictor más fuerte es la atribución de daño ($r = .524$), seguido por la edad ($r = .419$).

En el estímulo control, no hay correlación entre atribuir daño y moralizar el estímulo. Hay una correlación positiva pero pequeña de la moralización con la edad; ni el sexo ni el estrato son predictores significativos de la moralización del estímulo control.

Como último análisis presentamos una regresión lineal. La moralización de los actos impuros es la variable dependiente. La atribución de daño y las tres variables demográficas son los predictores y se introducen simultáneamente en el modelo. Las correlaciones de la Tabla 2, arriba, nos dicen que la atribución de daño predice significativamente, y con efecto fuerte, la moralización. Pero también muestra que la edad tiene un peso predictivo muy similar. Podría suceder, por ejemplo, que la edad prediga mejor, estadísticamente hablando, la moralización. La regresión nos ayuda a saber si esto es así, pues controla la confusión entre variables. Nos muestra la significancia estadística de los predictores descontando cualquier eventual confusión de efectos entre ellos y nos permite inferir, en caso de que el efecto de ambas variables estuviese confundido, cuál de las dos variables recoge mejor el efecto. Podría ser que la edad –y no la atribución de daño– sea el mejor predictor. La atribución de daño podría no ser, incluso, un predictor significativo.

Tabla 3. Coeficientes de la regresión lineal^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.193	.113		1.713	.089
	AtribDaño_pureza	.083	.014	.407	6.097	.000
	Edad	.009	.003	.223	3.273	.001
	Sexo	-.110	.050	-.140	-2.196	.030
	Estrato	-.059	.024	-.159	-2.491	.014

a. Dependent Variable: Moraliz_purez

La regresión nos muestra que todas las variables del modelo tienen un efecto significativo sobre la moralización. En particular, el coeficiente estandarizado Beta nos muestra que el efecto de la atribución de daño es casi el doble del efecto de la edad y casi el triple –positivo– del efecto negativo del sexo y del estrato (los varones y los individuos de estrato alto moralizan menos). El efecto de la atribución de daño es, estadísticamente, el más fuerte.

6. Discusión y conclusiones

Los resultados anteriores nos muestran que la atribución de daño es el predictor más fuerte de la moralización de los actos impuros que Haidt et al. (1993) clasificaron como “males sin daño”. Nuestros resultados fortalecen la crítica de Turiel a Shweder cuando el primero sostenía que probablemente las culturas occidentales difieren de las no occidentales en que no ven daño en algunas acciones tradicionalmente concebidas como impuras. Aquí se ve que dentro de una muestra de la población colombiana, la atribución de daño en las acciones impuras que involucran la sexualidad varía en función de la atribución de daño a otros con esas acciones, la edad, el estrato y el sexo. Estos resultados no nos dicen qué razones esgrimen las personas para pensar que las acciones impuras del estudio producen daño a otros. Podemos especular que los participantes piensan que el agente de esas historias daña su carácter y que, posteriormente, podría causar daño a otros en sus interacciones sociales. Este problema, siendo importante, es sin embargo secundario para los fines de este trabajo. Lo importante es que queda cuestionada la existencia de “males sin daño”.

Sin duda, Haidt y colaboradores podrían aun sostener que la conexión entre daño y moralización es una racionalización *post-hoc*. Las personas atribuyen daños a conductas que juzgan inmorales como una forma de justificar su juicio ante otros, no porque piensen en el daño como la razón de su juicio condenatorio. La verdadera razón quedaría oculta a su conciencia. Aunque esto es posible, es una hipótesis que tendría que estar motivada por alguna teoría que explicase por qué camuflamos con razones de daño las condenas morales que hacemos por otras razones y por qué esas razones quedan fuera de nuestra conciencia. Nuestro estudio no pretende demostrar que esta posibilidad está excluida. Pero es claro que resulta arbitrario sostenerla sin ninguna prueba.

Lo que nuestros resultados demuestran es que la pretensión de haber encontrado pruebas empíricas de la existencia de conductas malas que no se perciben como causando daño no tiene apoyo experimental hasta el momento. El debate está abierto sobre si el daño es la causa de la moralización o si ver una acción como inmoral es causa de que le atribuyamos *post-hoc* un efecto dañino sobre la comunidad o sociedad. Pero queda claro que los actos impuros están asociados al daño en la mente de los participantes colombianos de nuestro estudio. Para avanzar en el tema, los estudios experimentales necesitan ahora diseños más complejos; diseños que puedan discernir entre la racionalización, por un lado, y el razonamiento que nos lleva de la atribución de daño a la emisión de un juicio moral condenatorio, por otro. Tales avances están aún por venir.

Referencias

- Flanagan, O. (2017), *The Geography of Morals. Varieties of moral possibility*, Oxford University Press, Nueva York.
- Haidt, J. (2001), "The emotional Dog and its Rational Tail: A social intuitionist approach to moral judgment", *Psychological Review*, 108, 4: 814-834.
- Haidt, J. (2007), "The new Synthesis in Moral Psychology", *Science*, 316, 5827: 998.
- 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: 626.
- Kelly, D., Stish, S., Haley, K., Eng, S. & Fessler, D. (2007), "Harm, affect, and the moral/conventional distinction", *Mind & Language* 22(2), 117-131.
- Miller, J. G., Bersoff, D. M. & Harwood, R. L. (1990), "Perceptions of social responsibilities in India and the United States: Moral imperatives or personal decisions", *Journal of Personality and Social Psychology*, 58: 33-47.
- Nucci, L. (2001), *Education in the moral domain*, Cambridge University Press, Cambridge.
- Rosas, A. (2013), "Harm, Reciprocity and the Moral Domain", *Perspectives and Foundational Problems in Philosophy of Science*, 493-502.
- Shweder, R. A. (1990), "In defense of moral realism: Reply to Gabennesch", *Child Development*, 61: 2060-2067.
- Shweder, R. A. & Sullivan, M. A. (1993), "Cultural psychology: Who needs it?", *Annual Review of Psychology*, 44: 497-52.
- Shweder, R. A., Mahapatra, M. & Miller, J. (1987), "Culture and moral development", En J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children*, (pp. 1-83). University of Chicago Press, Chicago.
- Stich, S., Fessler, D. & Kelly, D. (2009), "On the morality of harm: A response to Sousa, Holbrook and Piazza", *Cognition*, 113, 93-97.
- Turiel, E. (1983): *The Development of social Knowledge: Morality and Convention*, Cambridge University Press, Cambridge.
- Turiel, E., Killen, M. & Helwig, C. (1998), "Morality: Its structure, functions, and vagaries", En J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 155-244). University of Chicago Press, Chicago.