

DOI: <https://doi.org/10.22201/iifs.18704905e.2024.1575>

UN RECORRIDO ANALÍTICO SOBRE LA CULPA DEL SOBREVIVIENTE

Carlos Muñoz-Serna, y Carlos G. Patarroyo Gutiérrez, *Culpa sin trasgresión. Un análisis filosófico de la culpa del sobreviviente*, Lambda, Ciudad de México, 2023, 152pp., ISBN–10: 6075971564, ISBN–13: 978–6075971568

JULIANA MEJÍA QUINTANA
Universidad del Rosario
Escuela de Ciencias Humanas
Colombia
julimequi@gmail.com
<https://orcid.org/0000-0002-1279-7227>

1. *Introducción*

El ser humano es sujeto de emociones y creencias, agente de sus decisiones, permeado por su entorno cultural; la forma como se articula la interioridad y exterioridad ha sido tema de amplio debate en la filosofía en cuestiones como la relación mente-cuerpo, la responsabilidad moral evaluada según la libertad o el determinismo, y el rastreo de intencionalidad como causa o no de las acciones morales.

El caso de la culpa del sobreviviente es un fenómeno que permite evaluar dichas relaciones para comprender un poco más lo constitutivo del agente moral. La culpa del sobreviviente se presenta en quienes, sabiendo que no tienen parte causal en la acción que lamentan, o que no habrían podido evitar, se afirman como culpables. Este sentimiento parece contradictorio desde una mirada cognitivista de las emociones. No obstante, la propuesta que aquí se evaluará corresponde a una defensa, desde la filosofía analítica, a teorías cognitivistas de las emociones, mostrando —con una tipología de la culpa, un análisis de caso, y una adaptación del modelo deflacionista del autoengaño— que este fenómeno no es contraejemplo a dicha postura, sino consecuencia de fundamentos sesgados de las creencias del sujeto.

El marco de análisis será la disputa que puede tener un agente moral entre sus juicios racionales, sus emociones e intenciones al enfrentarse a casos límite en los que una consecuencia inesperada o negativa lo invitan a evaluar la cadena causal de los acontecimientos y la subsecuente afectación de estos en su propia historia de vida y en la de quienes lo rodean. El libro *Culpa sin trasgresión. Un análisis filosófico de la culpa del sobreviviente*, nos sitúa ante casos de pérdidas y duelos desde los que emergen preguntas frente a la culpa y la responsabilidad moral. ¿Conviene en esos casos dejar dormir la razón? o ¿vale la pena evaluar y valorar lo que motiva nuestras creencias y sentimientos?, ¿estamos dispuestos a asumir lo que implica la respuesta a estas preguntas?

En este estudio crítico, mantendré el énfasis en los agentes morales, para afirmar con los autores del libro que los casos de la culpa del sobreviviente no son contraejemplos a la teoría cognitivista de las emociones. Sin embargo, tomaré distancia frente a la premisa sustancial de dicha tesis, ya que contrario a lo que se afirma en el libro, sugeriré que la culpa analizada en dichos casos, más que culpa por sobrevivir, responde a culpas asociadas a factores que sí están en control de los agentes, y que depende en gran medida del compromiso que tenga cada persona con la verdad.

2. Reconstrucción

Carlos Muñoz-Serna y Carlos G. Patarroyo Gutiérrez nos llevan paso a paso a analizar el fenómeno emocional de la culpa desde un ejercicio de filosofía analítica que da cuenta de la lógica desde la que se construyen y evalúan los estados mentales de un agente moral. Dicha evaluación, en el abordaje de las posturas cognitivistas, da cuenta de cómo un juicio racional que se espera que sea coherente en el sujeto que lo anuncia, —en el caso de la *culpa del sobreviviente*— parece ser un contraejemplo a esta implicación lógica, pues estamos ante un sujeto que afirma sentirse culpable de algo que sabe que no causó ni pudo evitar.

Contrario a ello, se sugiere que la culpa del sobreviviente no es un contraejemplo a las posturas cognitivistas, toda vez que no estamos ante un caso de creencias contradictorias en el sujeto, sino más bien de un sesgo en la formación de dichas creencias. La defensa de esta tesis implica comprender muy bien las condiciones de la culpa, el modelo cognitivista de las emociones, y la agencia moral desde la formación de creencias.

Para ello, se aborda tanto el modelo cognitivista de las emociones, como las condiciones de la culpa en general, con el fin de distinguirla de la culpa del sobreviviente. Además, se exponen cuatro casos puntuales que nos permiten evaluar las condiciones propuestas a la luz de situaciones cotidianas:

El primero, padres que pierden a sus hijos por enfermedades; el segundo, un joven que sobrevive a un incendio en el que pierde a su madre y a su hermana; el tercero, el fallecimiento de tres artistas del rock en un accidente aéreo, al que sobrevive Jennings luego de ceder su asiento a uno de ellos en el fatídico vuelo; el cuarto, el de Jacobo Timerman, periodista que sobrevive a un tortuoso secuestro durante la dictadura Argentina en los años setenta.

Posteriormente, y como respuesta a la objeción de ver la culpa del sobreviviente como contraejemplo a las tesis cognitivistas, se presenta el modelo deflacionista de Mele, que si bien está hecho para explicar una noción de autoengaño, sirve de guía para explicar la formación de creencias en el sujeto, y cómo estas pueden ser sesgadas sin necesidad de afirmar una escisión del sujeto en términos cognitivos. En este punto se enfatizan dos creencias puntuales: la creencia en un mundo justo, y el arraigo; dos creencias que, de abandonarse, implicarían un costo en la forma como el sujeto se entiende a sí mismo y como juzga la responsabilidad moral tanto individual como de terceros.

Desde el modelo cognitivista de las emociones,

las emociones no son algo aislado de los aspectos racionales de los sujetos [. . .], si los procesos cognitivos hacen también parte de nuestros procesos emocionales, es de esperarse que haya una cierta coherencia entre nuestras creencias, nuestros juicios, nuestras evaluaciones de las evidencias y nuestras emociones. (p. 77)

Situarse en la postura cognitivista implica rechazar que las emociones sean solo sensaciones que experimenta el sujeto, ya que esto impediría explicar cómo interactúa una emoción con otras actitudes y con la formación de creencias. Se debe refutar también que nuestras acciones y juicios solo se suceden de las emociones sin una relación causal; en lugar de ello,

las emociones no son solo causas de nuestras acciones, sino que son motivos o razones para estas. La relación entre emoción y acción no puede ser reducida a una mera relación entre dos eventos que se suceden. Las emociones dan sentido a la acción y funcionan en algunos casos como racionalizadoras de conductas particulares. (p. 83)

Esta postura responde a una fuerte crítica a la psicología conductual que:

tiene como base una noción darwinista de la conducta humana según la cual el comportamiento humano es, fundamentalmente, adaptativo (respecto a las circunstancias). ¿Dónde queda el comportamiento capaz de transformar críticamente las circunstancias? No lo hay, no se toma en cuenta. Así de simple. La terapia cognitiva enseña a adaptarse a la realidad social, sea ésta injusta o no. En una “distorsión cognitiva” el equivocado es el paciente, no el mundo o el futuro previsible. Se trata de adaptar al ser humano a las circunstancias, sin que éstas sean criticables. (Sáez Rueda 2020)

Contrario a ello, tachar de irracional al sujeto en el caso de la culpa del sobreviviente es no prestar atención a sus emociones y desacreditarlas sin siquiera analizar de dónde proceden o por qué las mantiene. De ahí que, para hacer énfasis en el sujeto y en su agencia, se adopte una noción de emociones —de la mano de Lyons (1993)—, estructuradas por juicios que dan lugar a creencias; evaluaciones objetivas o apreciaciones subjetivas que se relacionan con el subsiguiente estado mental de una emoción particular.

Lejos de una postura dualista, la tesis cognitiva a la que apelan los autores afirma la interdependencia entre la emoción y la cognición así:

la activación de una respuesta emocional está vinculada básicamente a los procesos de valoración. Si la valoración del evento es positiva o negativa se generará un tipo de patrón psicológico igualmente positivo o negativo que se evidenciará en una expresión característica de una emoción [...] una teoría cognitiva de la emoción es, por consiguiente, aquella que incluye algún aspecto del pensamiento, usualmente una creencia o un juicio, central para el concepto de emoción; de esta forma se aleja la emoción de la tendencia al irracionalismo. (Lyons 1993; citado en Pinedo Cantillo, Pacheco Castañeda, y Yáñez Canal 2017, p. 87)

Siguiendo esta postura cognitivista, diríamos que los casos de la culpa del sobreviviente apelan a momentos en los que parece haber un fallo en la forma como se procesa una emoción, o en cómo se interpreta la información al punto que se resiste a cualquier evidencia o no responde a las razones adecuadas; verbigracia, manteniendo un sentimiento de culpa por algo que no estaba en la línea de acción del agente.

Es por esta razón que la culpa del sobreviviente sería un contraejemplo, pues cuando la persona encuentra las evidencias que le eximen de la culpa por sobrevivir, debería cambiar su emoción de culpa, cosa que no sucede en los casos evaluados.

Si hemos afirmado que las emociones están vinculadas a creencias y juicios, entonces surge la posibilidad de la responsabilidad por nuestras emociones; si las emociones son juicios y se pueden debilitar o transformar por la introducción de otras creencias y otros juicios, es evidente que somos en buena parte responsables de ellas. (Pinedo Cantillo, Pacheco Castañeda, y Yáñez Canal 2017, p. 92)

Ahora bien, mantener el vínculo irracional-racional de modo tan tajante impide evaluar si quien permanece con un sentimiento de culpa en estos casos puede estar, más que ante una actitud irracional, ante una creencia sesgada que le impide relacionar las evidencias con la verdad.

Para el caso de la culpa, el juicio que la causa implica la conciencia de haber actuado de un modo que el sujeto evalúa como moralmente reprochable. Lo interesante de estos casos es que “la persona tiene evidencias que soportan la creencia de ser inocente o víctima en lo sucedido, por lo tanto, no debería derivar de allí la emoción de la culpa” (p. 11). En los casos propuestos, los sobrevivientes afirman sentir culpa de seguir con vida mientras sus familiares o amigos no lo están, y afirman, además —contra toda evidencia— sentirse culpables por el hecho mismo de que sus familiares o amigos hayan muerto. Si se tiene en cuenta que ambas creencias aparecen como contradictorias, la culpa del sobreviviente parece ser un contraargumento a la concepción cognitivista de la emoción. Pero esto es justamente lo que Muñoz y Patarroyo evaluarán en aras de explicar que no hay creencias contradictorias en el sujeto.

Para ello, el primer paso es evaluar las condiciones que hacen de la persona un agente moral, esto es, haber estado involucrado en lo acontecido, tener capacidad de actuar intencionalmente y poder ser elogiado o culpado como parte de la acción que se evalúa.

Teniendo en cuenta que el elogio o la culpa se atribuyen a un agente moral por la autoría de la acción, cabe destacar que esta autoría frente a las causas y consecuencias de una acción requiere de tres condiciones: a) el control guía,¹ b) la condición epistémica, y c) la respuesta a razones.

¹ En este punto se recoge una discusión de largo alcance en la filosofía analítica en relación con el Principio de Posibilidades Alternativas (PPA) que parte de

En la condición (a) “decir que X acto lo ha hecho él supone a su vez que tenía algún control sobre su ocurrencia” (p. 36). La condición (b) implica ser consciente de la valoración moral que realiza sobre su acción: “significa que tiene una serie de creencias (no solo conocimientos) acerca del valor moral de sus acciones o de las consecuencias y, a sabiendas de este carácter, elige o no realizar la acción” (p. 37); la condición (c) implica que ese control (condición a) que el agente tiene de sus acciones, es tal (en términos de conciencia de la misma: condición b) que puede modificar su conducta dependiendo de los escenarios en los que se encuentre; esto hace más fácil decir de esa acción que fue *suya* (Muñoz-Serna y Patarroyo Gutiérrez 2023).

Una vez establecido en qué medida un agente moral es autor de la acción, los autores generan la siguiente tipología frente al concepto de culpa:

- *Culpa-G (Guilt)*: Strawson la define como “la actitud reactiva, reflexiva o autorreactiva que un agente siente respecto de su propia acción cuando ha producido un daño o mal a alguien” (p. 43).
- *Culpa-ACC (Accountability)*: Rendición de cuentas: “hay una legítima demanda de una parte a la otra por algo. Cuando esta demanda es violada o incumplida, entonces una de las partes (aquella a la que se ha incumplido) tiene el derecho de exigir, de reclamar a la otra parte” (p. 45).
- *Culpa-B (Blame)*: Asignar una sanción a quien le atribuyo la categoría de culpable.
- *Culpa-ATT (Attributability)*: El juicio mental de que alguien actúa mal, “pero no nos es permitido acompañar este juicio con la posterior expresión de una actitud reactiva” (p. 47).

Estas distinciones permiten ver que los juicios sustentados en ciertas demandas morales frente a las acciones se dirigen tanto hacia los demás como hacia nosotros mismos, y de allí se siguen determinados sentimientos o emociones. Para la emoción concreta de la culpa, “la

entender la libertad como tener posibilidades de acción. Frente a este principio hay quienes afirman que dichas alternativas deben ser reales y absolutas para que exista libertad y, por tanto, responsabilidad moral, y quienes niegan la necesidad de que dichas alternativas sean reales y absolutas; más bien deben evaluarse las alternativas causalmente relevantes en la acción que está en control del sujeto (este sería el caso en los ejemplos de culpa del sobreviviente).

emoción sucede a la acción supuestamente transgresora, ella aparece porque el sujeto se vuelve consciente de lo que ha hecho y encuentra que ha dañado algo que valoraba, ha lastimado a alguien” (p. 55).

El punto central aquí es ver que, en ninguno de los casos de la culpa del sobreviviente, el sentimiento se puede explicar exhaustivamente desde alguna de las cuatro definiciones dadas anteriormente. Lo que vemos en el caso de la culpa del sobreviviente es un tipo de Culpa-G en tanto actitud reflexiva o autorreactiva “conectada con el deseo de que la acción que constituye la ofensa nunca hubiese sucedido” (p. 57), pero no está dada con base en el juicio de una acción en la que el agente haya tenido las tres condiciones ya expuestas para la atribución de la acción como suya.

En ninguno de los casos estudiados el agente tenía control guía o condición epistémica; ningún sobreviviente “tenía motivo, creencia o evidencia que llevara a pensar que actuar de manera diferente habría evitado la tragedia [...] ninguno de los casos presenta una situación en la cual el resultado de la misma fuera deseado por el agente” (p. 69); esto lleva a concluir que el fenómeno de la culpa del sobreviviente no encaja en ninguna de las tipologías de la culpa, ya que si no se cumplen las condiciones necesarias para la Culpa-ACC y para la Culpa-ATT, no se daría la Culpa-B ni la Culpa-G que son subconjuntos de la primera. En ese sentido, y siendo la culpa del sobreviviente un caso distinto a las cuatro tipologías anteriores es importante evaluar por qué el agente afirma sentir Culpa-G.

La propuesta para trabajar esta paradoja es muy interesante, y da lugar al análisis del modelo deflacionista de Mele frente al autoengaño, y la formación de creencias, específicamente la creencia en un mundo justo, y la fuerza del arraigo.

La razón de acudir al modelo deflacionista de Mele es que, a diferencia de lo que ocurre con las posturas tradicionales que afirman que el autoengaño es posible asumiendo que un agente puede sostener simultáneamente dos creencias contradictorias, este “no pone en peligro los estados mentales del sujeto” (pp. 91–92). Las posturas tradicionales del autoengaño sugieren

rechazar la idea de un yo unificado y defender más bien la idea de un yo compartimentado. De esta manera, se podrían ubicar las creencias contradictorias entre sí en compartimentos diferentes de manera que ambas pudieran estar presentes sin que por ello entraran en conflicto abiertamente. Pero en este caso el problema está en cómo se puede cumplir con la condición según la cual el sujeto ha de reconocer de alguna manera la contradicción de sus creencias. (p. 96)

Contrario a la tesis anterior, si “la compartimentalización es tal que no se pueda dar este reconocimiento, se pierde el autoengaño pues, en palabras de Okseberg Rorty ‘se ha abandonado la identidad del engañador y el engañado’” (Rorty 1988, p. 22; citado en Muñoz-Serna y Patarroyo Gutiérrez 2023, p. 97). Es así como volvemos a la tesis de Mele según la cual, en lugar de tener creencias opuestas o compartimentadas, lo que existe es una motivación sesgada en alguna de esas creencias: “El agente tiene un fuerte motivo para creer que P, y este motivo lo lleva a tratar sesgadamente, de manera no intencional, la evidencia de la que dispone. Y este tratamiento sesgado es lo que lleva a creer que P cuando en realidad, la evidencia favorece no-P” (p. 99). Más allá de si se avala esta postura frente al concepto de autoengaño, lo que interesa es la posibilidad de formar creencias que resisten a la evidencia.

En los casos de la culpa del sobreviviente, existen dos creencias específicas como causantes de dicho sesgo en la creencia de culpabilidad. A saber: la creencia en un mundo justo y el arraigo. Veamos cada una.

Existe una importante relación entre la creencia en un mundo justo y la creencia en un sistema de responsabilidad moral desde dos componentes: “por un lado, el orden, esa idea de que todo lo que ocurre sucede por una causa explicable y rastreable. Por el otro, en el ámbito de las relaciones humanas, la idea de que cada quien recibe lo que merece” (p. 114); es por esto que dichas creencias sustentan emociones reactivas como la indignación.

Como segunda creencia estaría el arraigo que podemos tener hacia alguien o algo, y cómo esto nos hace vulnerables:

podemos reconocer el valor de algo, sin por ello valorarlo [...]. Lo importante aquí reside en dos puntos esenciales: en primer lugar, que no estar arraigado a algo no implica desconocer su valor y, en segundo lugar, que el arraigo conlleva reacciones emocionales y afectivas que no solemos tener respecto de lo que acontece con personas o situaciones frente a las cuales no tenemos arraigo alguno. (p. 117)

Cuando en el sujeto se unen la creencia *rigurosa* en el mundo justo y un arraigo *fuerte* respecto de cierto objeto o persona, entonces percibe un alto costo en renunciar a esa creencia, y esto sesga las evidencias que puede tener frente a hechos que pongan en duda dicho arraigo o dicha creencia en el mundo justo.

Con esta última tesis, los autores cuentan ya con las piezas completas del rompecabezas, y describen la culpa del sobreviviente como

un fenómeno que entraña un juicio de responsabilidad reflexivo. Es decir, “el individuo que presenta esta emoción tiene una forma de Culpa-G que a su vez está estructurada por un juicio de Culpa-ACC que dirige a sí mismo” (p. 125). Es clave este juicio autorreflexivo toda vez que, en este caso, permea un juicio reflexivo como lo es el del mundo justo, y en lugar de mantenerlo como un juicio externo,

las personas parecen usarlo hacia sí mismos, esto deriva en la búsqueda de *escenarios alternativos* frente a lo que sucedió; si bien esta búsqueda es un ejercicio más imaginativo, es la herramienta sesgada del sujeto para buscar *control* de la situación, no frente a lo que ya está saldado, sino frente a lo que aún puede hacer. (p. 135)

Así, el sujeto deja de verse como víctima, para verse como agente de la acción, y es por eso por lo que puede terminar sintiendo culpa. El punto clave con esta conclusión es que no estamos ante creencias irracionales, sino ante un tratamiento sesgado de las mismas por la motivación ya expuesta.

3. *Evaluando las piezas del rompecabezas*

El texto nos conduce poco a poco por diversos debates de la filosofía sintetizando lo pertinente para el análisis del problema de la culpa del sobreviviente; esto exige del lector ir llevando el hilo argumentativo, las *piezas de rompecabezas*. Lo que me propongo en este segundo apartado es cuestionar el abordaje de algunas de esas piezas, analizando elementos que —por no ser el eje central del libro— se pasan con rapidez, pero que cuestionarlos permitirá volver críticamente sobre la conclusión propuesta.

El Principio de Posibilidades Alternativas – PPA

Un elemento central para la atribución de culpa es que la acción juzgada pueda atribuirse a un agente moral. Esa atribución depende del control que tiene el agente sobre la acción. En este punto, los autores descartan rápidamente la necesidad del control regulativo, para quedarse con el control guía; este último tampoco estaba presente en la causa de muerte de los familiares o amigos del sobreviviente. Sin embargo, me parece clave evaluar aquellas acciones u omisiones sobre las que sí tenían cierto control, y que se enfocan más en cómo el sujeto se entiende frente a aquel con quien tiene un vínculo afectivo que genera unas obligaciones morales tácitas.

En el caso de los padres que pierden a sus hijos por enfermedad, una de las razones que aducen es que “se ven a sí mismos como parcialmente responsables de lo sucedido, en tanto que no pudieron detectar los síntomas de la enfermedad con anterioridad” (p. 18). Aquí se alude a un resquicio mínimo de responsabilidad: cuestionan si su tardanza en la identificación de la enfermedad pudo generar un cambio en la cadena causal, si bien las explicaciones médicas descartan que actuar a tiempo hubiera podido cambiar la situación, el juicio está sobre la noción de cuidado con que se juzgan los padres, y es así de exigente justamente por tratarse de un vínculo afectivo fuerte.

En el mismo caso, hay quienes se culpan “por haber sido demasiado severos con sus hijos, o no dedicarles más atención” (p. 19). Nuevamente estamos ante un escenario que no está relacionado causalmente con la enfermedad, pero sí con las obligaciones morales que evalúan como incumplidas en relación con la víctima. En ese orden de ideas, si bien la afirmación generalizada es que los padres se sienten culpables *por sobrevivir a sus hijos*, es una expresión que atañe a causas más pequeñas y concretas como un bajo cuidado preventivo, falta de tiempo invertido con ellos, o por la forma como llevaron la situación (rabia, desesperación...), es decir, eventos que sí están en su control.

El hecho de afirmar la culpa como relacionada con la causa de muerte, puede tener que ver con que ese es el resultado que salta a la vista más evidentemente y del que se tienen que ocupar; sin embargo, los sobrevivientes en ningún momento se enuncian como agentes intencionales de la muerte de sus familiares o seres queridos.

En el segundo caso estudiado, Carlos es un joven que se siente culpable por sobrevivir a su madre y a su hermana luego de un incendio. En este caso, encontramos también juicios bajo los cuales Carlos se lee a sí mismo como un *mal hijo*, siente que no cumplió cabalmente con las obligaciones morales de hijo mientras su madre estaba viva; dado el tipo de vida que llevaba, se sentía con menos mérito de sobrevivir a su hermana y a su madre. Nuevamente, más que una culpa por la muerte de sus familiares, la cual claramente no causó, la culpa por sobrevivir está atada a otra serie de obligaciones morales incumplidas en su rol de hijo, aspectos frente a los que sí se sabe personalmente responsable.

Lo que quiero mostrar con estos ejemplos es que lo que resienten los sobrevivientes no es estrictamente pensar que son parte de la cadena causal de la muerte o daño sufrido por sus seres queridos;

sino que, dichas muertes o daños, al ser traumáticos desatan sentimientos que obstaculizan la evaluación causal real de la culpa; por eso sitúan el sentimiento de culpa en la consecuencia: muerte de uno y supervivencia del otro. Sin embargo, es un sentimiento que está basado en eventos puntuales sobre los que sí tenían cierto control, o sobre los que juzgan que podrían haber actuado diferente. En ese orden de ideas, sigue manteniéndose como necesaria para la culpa el Principio de Posibilidades Alternativas.

La creencia en un mundo justo

La primera afirmación que merece ser evaluada es: “parece ser evidente que el mundo es injusto” (p. 114). Considero que no es algo que pueda darse por sentado; una cosa es que existan innumerables injusticias, y otra que el mundo sea en sí mismo algo injusto; de lo primero no se sigue necesariamente lo segundo. La importancia de evaluar esta hipótesis radica en que constituye una de las creencias principales que parecen llevar a un agente a evaluarse dentro de un fenómeno como lo es la culpa del sobreviviente.

Los autores acuden a los estudios empíricos realizados por Faccenda y Pantaleón (2011), en cuyos resultados se demuestra una relación entre tener una creencia arraigada en el mundo justo, y cómo esta creencia lleva a evaluar y clasificar las situaciones de injusticia de un modo sesgado. Puntualmente, usando dicha creencia como justificación de ciertas injusticias. No obstante, considero que la respuesta de los participantes que afirman que la víctima *se lo merecía*, puede leerse más como un caso de desconexión moral (Bandura 1999) que como una creencia en un mundo justo.

El psicólogo Albert Bandura afirma que la agencia moral puede irse empobreciendo con base en creencias y prácticas que el sujeto va normalizando y que lo llevan a lograr separarse de una autorregulación moral frente a sus acciones. En la desconexión moral, el sujeto cae en prejuicios, en afirmaciones generalizadas, o en atribuirle la culpa a la víctima, entre otros mecanismos que tergiversan la evaluación de su acción al punto de sentirse ajeno a ella (cambia la noción de agencia) y, en consecuencia, no se siente comprometido con su acción. Este parece ser el caso de los participantes del experimento, que son ajenos a la situación que analizan y carecen del contexto suficiente para evaluarla.

Si de entrada se entiende la creencia en el mundo justo como negativa, se deja de lado la posibilidad de usar dicha creencia como una motivación crítica del sujeto para trabajar en el ideal de justicia, por

ejemplo, fomentando la empatía para la búsqueda de la justicia en el mundo. Es así como la afirmación de dicha creencia sin más (tanto en el mundo justo, como en la imposibilidad de este), puede ser un mecanismo de desconexión moral que sirve de justificación al agente para extrapolar su juicio hacia “el mundo”, más que hacia su acción en ese mundo concreto.

4. *Motivaciones sesgadas*

Hay un punto en el modelo de autoengaño sugerido por Mele y adoptado por los autores, que no me parece del todo claro, a saber, que el sesgo en la motivación es un hecho no intencional.

Es importante entender en dónde se enmarca la afirmación de Mele frente al debate sobre la naturaleza del autoengaño. En este debate hay quienes evalúan el fenómeno en paralelo con el engaño que se genera a nivel interpersonal, y quienes se distancian de evaluar este fenómeno desde esa comparación. Adicionalmente, hay dos posturas entre quienes afirman que este fenómeno es posible y que somos responsables del autoengaño; están allí los intencionalistas que defienden la necesidad de la intención de engañar en el sujeto, y los no intencionalistas como Mele, que defienden que, al no tratarse de un caso igual al del engaño interpersonal, no se requiere la intención de engañar en el sujeto.

Con la postura no-intencionalista a la que se suman los autores del libro aquí analizado, buscan salir de un problema que presenta la postura intencionalista, a saber, que la intención de engañar está de algún modo oculta al agente, elemento que hace perder el control necesario para la atribución de responsabilidad, o bien que —como defiende Davidson (1985)—, las creencias contradictorias en el sujeto son posibles gracias a una partición de la mente, es decir, a un yo compartimentado. Para Mele, ninguna de estas dos aproximaciones es necesaria para explicar el autoengaño ya que lo importante es la motivación que genera la creencia falsa, no por eso contradictoria. En la motivación es donde se enmarca el control que es el elemento necesario para la responsabilidad. Así lo explica Coll:

En muchos casos de autoengaño lo que ocurre es, según Mele, que el deseo de que *p* provoca que se adquiera la creencia de que *p* y la no adquisición de no-*p*, sin necesidad de atribuir conocimiento previo de la verdad, ni atribuir a dicho conocimiento un papel causal [...] las emociones son suficientes para predisponer a alguien a adquirir una creencia sin necesidad de suponer, además, que es el conocimiento de la verdad del asunto el que provoca el autoengaño [...] Todo lo que

se requiere en tales casos es que no se sepa que p, se desee que p, que dicho deseo lleve al agente a tratar la información de manera distorsionada y que adquiera finalmente la creencia de que p. Ningún otro componente afectivo ni cognoscitivo es para Mele condición necesaria para explicar su estado. (2007, p. 151)

Como vemos, en el modelo de Mele, se defiende la agencia de la persona, se da énfasis en la emoción y no se habla de una contradicción cognitiva en el agente. Dichas razones son las que llevan a Muñoz-Serna y Patarroyo a adoptar este modelo. Sin embargo, mi pregunta es, ¿por qué llamar a esto autoengaño en lugar de simplemente llamarlo error?, y ¿en qué medida no estar comprometidos con la verdad afecta la noción de agencia moral?

La respuesta de Mele está en el control que tenemos sobre esas creencias sesgadas, pero nuevamente esto afirmarí que, en el caso de la culpa del sobreviviente, si lo que se evalúa es aquello sobre lo que tienen control, los casos no son de culpa por sobrevivir sino de aquellas creencias que sí estaban en sus manos. El punto decisivo a mi juicio está en la afirmación de Mele bajo la cual no es el conocimiento de la verdad lo que provoca el engaño. Como mostraré en el siguiente apartado, es justamente este punto el que nos permitiría hablar de una coherencia cognitiva en el sujeto para la imputación de responsabilidad.

Considero que el hecho de que no sea intencional pone en juego uno de los requisitos de la culpa en tanto la intención es importante para atribuirle al agente un juicio sobre su acción, o en este caso, sobre la valoración que hace de la misma.

Mele afirma además que “es justamente la presencia de este componente motivacional lo que produce el fenómeno del autoengaño y lo que permite que, ante la misma evidencia, la mayoría de la gente forme la creencia de que P, cuando el sujeto con la motivación específica forma la creencia de que no-P” (p. 102). Esta diferencia entre cómo un externo forma la creencia, y cómo el sujeto implicado forma la creencia, no se distingue en la definición de Mele.

Cuando Mele afirma que hay una creencia errada, lo que busca es explicar por qué el error no se evidencia como una contradicción en el sujeto, pero cabría pensar que, en los casos de la culpa del sobreviviente más que una motivación sesgada sobre una de las creencias (la posibilidad de haber cambiado el rumbo de los hechos), lo que está sesgado es la lectura de la causa real de la culpa.

Al afirmar esto no se está ignorando el sentimiento del sujeto, al contrario, verlo así permitiría acompañar la evaluación de motiva-

ciones y creencias sobre el control del agente ante los hechos para lograr moderar así el sentimiento. En ese orden de ideas, es cierto que juicio y sentimiento van de la mano (tesis cognitivista), pero por la misma razón, la expresión de culpa por parte del sobreviviente en los casos expuestos es una afirmación imprecisa.

En otras palabras, diría que los casos presentados como “culpa del sobreviviente” no son contraejemplo a la teoría cognitivista de la culpa, porque no son en realidad casos de culpa por sobrevivir, sino casos de culpas asociadas al contexto de los sobrevivientes, pero culpas que al final, si se miran con lupa, terminan estando asociadas a creencias cognitivas válidas en las que sí había control o cierta intencionalidad.

Entonces, ¿por qué se genera el fenómeno de la culpa del sobreviviente? Por un lado, puede deberse a que se pongan primero los juicios y creencias que acompañan la necesidad de asumir la pérdida para poder seguir viviendo la cruda realidad que ahora afrontan a futuro, antes que los juicios evaluativos sobre las acciones pasadas.

El sentimiento de culpa del sobreviviente tiene más que ver con la lectura de sí mismo hacia alguna forma de relación anterior con los involucrados, que con una culpa como responsabilidad causal de sí mismo frente al acto en cuestión. En ese orden de ideas, la creencia puede tejerse de la siguiente manera: 1. Fui mal hijo, 2. Mi madre ya no está, 3. No puedo resarcir mi mala acción como hijo; esto último es lo que genera el sentimiento de culpa tras la muerte, no es un sentirse responsable causalmente de algún modo de esa muerte.

¿Pueden cambiarse los juicios sin apelar a la verdad?

Finalmente, ya vimos de la mano de las teorías cognitivas que las emociones están vinculadas a creencias y juicios. Si bien los autores denotan que el problema en los casos de la culpa del sobreviviente está en mantener la creencia fuerte en el mundo justo y el arraigo. Considero que más allá de esas dos posibilidades subyace el problema de si el agente está o no comprometido con la verdad. Aunque el cambio de creencia no se dé en respuesta a la evidencia, esto no se debe necesariamente a una falla cognitiva, por lo cual no estamos ante un contraejemplo al cognitvismo. En este caso, el sujeto no tiene la intención de evaluar la evidencia, solo de mantenerse en las premisas, juicios o creencias que le permitan al sujeto mantener su identidad, que le permitan no *perder la cara*.²

² Con *perder la cara* hago referencia a la tesis de Goffman (1970), en *Ritual de la interacción*, donde se defiende la necesidad de configurar una identidad para

En un contexto social la persona desarrolla creencias en relación con su contexto intersubjetivo. Y esas creencias organizan un espacio «interno» de juego en la medida en que pueden ser o no coherentes entre sí. Cuando las creencias de un individuo chocan entre sí se produce una *disonancia cognitiva* que se intenta eliminar, [...] muchas veces «no son los hechos, sino lo que pensamos sobre los hechos, lo que nos perturba». (Sáez Rueda 2020)

Estoy con Muñoz-Serna y Patarroyo en afirmar que en lugar de tratarse de un problema cognitivo en el sujeto, estamos ante un error del cual el sujeto no es capaz de salir solo mediante la evidencia que se le presenta; el énfasis que quiero añadir está en que mantenerse en el error en estos casos se debe a un relativismo que vuelve al sujeto ajeno a razones: si cualquier evidencia cae en ser solo una perspectiva que no puede evaluarse como verdadera de modo absoluto, no es útil para sacar a una persona del error. Desde un juicio relativista, se nubla la subsecuente evaluación de las emociones.

Aquí estoy con la crítica que hace Coll Mármol al modelo deflacionista de Mele ya que, para Coll, “la atribución de responsabilidad a quien se engaña presupone el conocimiento previo de la verdad por parte de éste” (2007, p. 145). Esto nos lleva a afirmar que sin verdad no hay responsabilidad:

si no se fuera responsable de dicho estado, si no se tuviera control sobre el mismo, la racionalidad de la creencia se vería seriamente afectada pues, ante cualquier creencia que uno pudiera tener, nunca se estaría seguro de si se trataba de una creencia que perseguía la senda de la verdad o más bien se trataba de un producto de procesos psíquicos autónomos sobre los que el sujeto no tiene control. (Coll Mármol 2007, p. 147)

Alfred Mele no solo renuncia a la intención de engañarse para explicar el autoengaño, sino también al conocimiento de la verdad de quien se engaña. ¿Cómo puede entonces seguirse llamando engaño? Saberse engañado en algún momento implica reconocer que la evidencia que se daba por sentado en determinado momento estaba sesgada, y el cambio de valoración en el agente es lo que da cuenta del control y la responsabilidad.

entrar en interacción con otros. La noción de sí mismo no se da de modo asilado a lo que el entorno espera del rol que se ejerce en cada caso.

Puede que la mente cuente con mecanismos que se activen de un modo no intencionado para manipular la información con la que se cuenta, pero es precisamente el que hayamos estado expuestos a esa información y el que hayamos reconocido que la verdad iba por otros derroteros lo que nos convierte en responsables de nuestro estado. (Coll Mármol 2007, p. 156)

Contrario a esto, en Mele no se requiere que el agente reconozca la verdad, así,

Mele olvida que el autoengaño no es sólo un estado, sino un proceso en el que siempre existe una tensión entre el reconocimiento de que la evidencia apunta hacia cierto lugar y nuestro deseo de que no sea así. El autoengaño tiene un punto final en el que la persona en cuestión termina creyendo sinceramente una falsedad que no tiene por qué convivir con el reconocimiento previo de la verdad contraria. [...] la clave sigue estando en un reconocimiento de la verdad que es el que ha generado los procesos distorsionadores de la misma. (Coll Mármol 2007, p. 157)

Sin el compromiso con la verdad, el agente, más que en autoengaño, está en un error del que no tiene cómo salir porque su intención no apunta hacia el reconocimiento de una evidencia como verdadera o falsa, sino más bien hacia el mantenimiento de su *cara*, hacia una identidad con la que pueda convivir.

5. Conclusiones

Sin duda alguna, el libro de Muñoz-Serna y Patarroyo Gutiérrez es un excelente ejemplo de un ejercicio de filosofía analítica, de lo que Deleuze y Guattari definen como filosofía: “el arte de formar, inventar, fabricar conceptos” (1993 [2001], p. 8). Un trabajo que permite clarificar el fenómeno de la culpa a partir tanto de casos cotidianos como de un análisis lógico de las creencias, juicios y emociones del agente moral. El trabajo detallado de análisis tanto de enfoques de investigación en psicología, filosofía, sociología, como del abordaje complejo de un concepto tan complejo como el de la culpa, hacen de este libro un aporte importante a la filosofía de la acción.

Llama la atención lo importante que se hace para el sujeto *mantener la cara* (Goffman 1970), seguirse percibiendo como un ser coherente tanto hacia sí mismo como ante los demás; frente a esta necesidad resulta muy acertada la propuesta de los autores de desviar el análisis hacia las motivaciones de los agentes frente a su formación de creencias, más que afirmar de entrada que están cayendo en una

contradicción; así, se le da prioridad al punto de vista del sujeto, antes que juzgarlo desde una teoría ya establecida.

Así mismo, un punto clave es mostrar cómo la formación de creencias está motivada por cómo percibimos el mundo (ej. la creencia en el mundo justo), y cómo valoramos las relaciones con los demás (arraigo); esto permite enfatizar que la culpa, si bien se dirige hacia sí mismo, no solo surge de una valoración autorreferencial, sino que también implica una relación yo-otro, o hacia sí mismo como otro.

Un punto que se echa en falta es una caracterización más completa de los involucrados, la situación previa, las consecuencias; elementos que no den lugar a tantas interpretaciones puntuales de los casos presentados.

A partir de este estudio crítico, quise sugerir un análisis más detallado frente a la formación de creencias morales en el sujeto, y los mecanismos que permiten mantenerlas o transformarlas de manera crítica, empática y responsable. Una pregunta clave en ese sentido es la pregunta por la posibilidad o no del sujeto fragmentado, y la importancia de la coherencia para la identidad. Una coherencia que está dada en relación con los otros, pues la identidad también parece ponerse en juego cuando una situación como la vivida por los sobrevivientes impide el ejercicio normal del reconocimiento en sentido de algo duradero; por eso la culpa también es duradera.

Un eje de análisis que podría profundizarse a partir de este trabajo sería la relación entre el *mantenimiento de la cara*, y la necesidad de coherencia para el mutuo ejercicio del reconocimiento. *Mantener la cara* hace referencia a

reacciones —muchas veces inconscientes— como las miradas, los silencios, o los gestos, elementos que llenan de sentido nuestras interacciones cotidianas, las relaciones personales que se dan entre sujetos, y todos los hechos que nos hacen asumir determinados roles frente a otros y seguir ciertas pautas de conducta para mantener la interacción andando. Mantener la cara sería ese ejercicio sin el cual la interacción no sería posible. (Mejía Quintana 2016, p. 54)

La interacción social y la configuración de la propia identidad presuponen la idea de una coherencia; coherencia que se ve afectada en aquellos casos en los que sí se presenta una contradicción de creencias, juicios o enunciados, y que se agrava cuando el sujeto no ve necesidad de evaluar la contradicción porque no presupone la necesidad de la coherencia para la identidad. En ese orden de ideas, es determinante pensar en la importancia de la verdad en la agencia

moral, la interacción social y la identidad, ya que, sin este horizonte de referencia, se obstaculiza todo ejercicio de reflexión filosófica

El aporte que se ofrece en el libro y en este estudio crítico, desde miradas distintas pero complementarias, permitirá seguir ahondando en el amplio campo de la filosofía moral, la filosofía de la acción, las teorías del interaccionismo, de la identidad narrativa y el reconocimiento, evaluando no solo casos puntuales, sino también los marcos de referencia desde los cuales se entiende a alguien como agente.

BIBLIOGRAFÍA

- Bandura, Albert, 1999, “Moral Disengagement in the Perpetration of Inhumanities”, *Personality and Social Psychology Review*, vol. 3, no. 3, pp. 193–209.
- Coll Mármol, Jesús Antonio, 2007, “Autoengaño y responsabilidad”, *Teorema*, vol. 26, no. 3, pp. 145–159.
- Davidson, Donald, 1985, “Incoherence and Irrationality”, *Dialectica*, vol. 39, no. 4, pp. 345–354.
- Deleuze, Gilles, y Félix Guattari, 1993 [2001], *¿Qué es la filosofía?*, Anagrama, Barcelona.
- Frankfurt, Harry, 1969, “Alternate Possibilities and Moral Responsibility”, *The Journal of Philosophy*, vol. 66, no. 23, pp. 929–939.
- Goffman, Erving, 1970, *Ritual de la interacción*, Tiempo, Buenos Aires.
- Goffman, Erving, 1959, *La presentación de la persona en la vida cotidiana*, Amorrortu, Buenos Aires.
- Lionel Faccenda y Nathalie Pantaléon, 2011, “Analysis of the Relationships between Sensitivity to Injustice, Principles of Justice and Belief in a Just World”, *Journal of Moral Education*, vol. 40, no. 4, pp. 491–511.
- Lyons, William, 1993, *Emoción*, Anthropos, Barcelona.
- Mejía Quintana, Juliana, 2016, *Pedagogías lúdicas para la paz: una mirada a las percepciones de violencia en niñas y niños de Ciudad Bolívar*, Universidad del Rosario, Bogotá.
- Muñoz-Serna, Carlos, y Carlos G. Patarroyo Gutiérrez, 2003, *Culpa sin trasgresión. Un análisis filosófico de la culpa del sobreviviente*, Lambda, Ciudad de México.
- Pinedo Cantillo, Iván Alfonso, Laura Pacheco Castañeda, y Jaime Yáñez Canal, 2017, “Las emociones y la moral: claves de interpretación desde una aproximación cognitiva”, *Tesis Psicológica*, vol. 12, no. 1, pp. 82–104.
- Rorty, Amélie Oksemerberk, 1988, “The Deceptive Self: Liars, Layers, and Lair’s”, en Brian P. McLaughlin y Amélie Oksemerberk Rorty (eds.), *Perspectives on Self-Deception*, University of California Press, Berkeley, pp. 11–28.

Sáez Rueda, Luis, 04 de mayo de 2020, *Crítica a la psicología cognitiva*.
Disponible en
<https://www.ugr.es/~lsaez/blog/entradas/2020/cognitivismo.htm>

Recibido el 13 de octubre de 2023; aceptado el 8 de marzo de 2024.