
SOBRE EL CONCEPTO POPULAR
DE ACCIÓN INTENCIONAL.
UNA REVISIÓN A LOS
ACERCAMIENTOS EXPERIMENTALES

DAVID FAJARDO-CHICA
SERGIO MONARES
GIMENA VILLEGAS
ANDRÉS ZULES

ABSTRACT. We review here some efforts made by social psychology and experimental philosophy to elucidate how does the concept of intentionality apply when collateral effects occur. We attend to some of the tests made and the hypotheses offered, as well as an overview of the discussions involved.

KEY WORDS. Intentionality, experimental philosophy, common sense, popular psychology.

0§. INTRODUCCIÓN

La filosofía experimental puede describirse de manera amplia como un movimiento dentro de la filosofía analítica, cuya principal premisa es que la apelación a intuiciones implica un compromiso empírico y como tal es susceptible de ser puesto a prueba y a su vez confirmado o rechazado. El compromiso empírico que se adquiere cuando una filósofa señala: “es intuitivo que...” o “desde el sentido común se acepta que...” o “es por todos entendido que...”, es justamente que tal cosa se dé, es decir, que tales intuiciones sean en verdad compartidas por la mayoría de personas. Caracterizar el sentido común no es algo que pueda hacerse desde el sentido común mismo ya que, al igual que en otros asuntos, este podría no ser una buena guía. Por ello, es importante hacer una evaluación sistemática del sentido común. En este artículo vamos a dar por sentada la importancia que los acercamientos empíricos al sentido común tienen para ciertos debates filosóficos. Discusiones más amplias sobre este punto

Departamento de Filosofía, Universidad del Valle, Medellín, Colombia.
david.fajardo@gmail.com

se pueden encontrar en Fajardo-Chica (2011), Nichols (2004), Knobe & Nichols (2008), Nadelhoffer & Nahmias (2007).

La relevancia del estudio empírico sobre el tema que nos ocupa parece aún más justificada. Si aceptamos que los conceptos de la psicología popular, como lo es el de acción intencional, son conceptos del sentido común, ¿qué mejor manera de estudiarlos que prestar atención a cómo efectivamente son usados desde el sentido común? Cabe resaltar en este momento la distinción entre una pregunta normativa acerca de *qué debe contar* como una acción intencional y la pregunta descriptiva, que aquí nos ocupa, acerca de *comúnmente cómo* se aplica dicho concepto.

En §1 haremos algunas consideraciones generales del panorama en el que se enmarca la discusión que trataremos en detalle. La psicología popular, como se ha entendido tradicionalmente, ha sido objeto de revisión desde varias perspectivas; señalaremos algunos aspectos de estas revisiones y cómo la discusión acerca del concepto de acción intencional encaja dentro de esta imagen. En §2 presentaremos los resultados de Joshua Knobe (2003) que dieron inicio al debate. Los resultados de Knobe muestran una asimetría en la aplicación del concepto de acción intencional respecto a efectos colaterales. Tal fenómeno se conoce como el *efecto Knobe* o el *efecto del efecto colateral*. Para explicar tal asimetría, Knobe concluye que las personas tienden a calificar como intencional un efecto colateral cuando es moralmente reprochable, mientras que si el efecto colateral de una acción es moralmente bueno, las personas tienen a pensar en él como un efecto no intencional. Revisaremos el experimento y expondremos por qué la solución ofrecida por Knobe suscita tanta polémica. En §3, comenzaremos la revisión de varias hipótesis explicativas, iniciando con un mapa de la discusión que facilitará la comprensión del debate. La discusión sobre cuál es la mejor explicación para el efecto Knobe ha sido extensa. En este artículo revisaremos ocho hipótesis explicativas distintas, mostraremos la evidencia empírica que algunas de ellas presenta y haremos una evaluación de sus bondades y defectos. Finalizaremos en §4 ofreciendo una evaluación del debate, conclusiones y unas rutas experimentales para el futuro.

1§. LA PSICOLOGÍA POPULAR A REVISIÓN

Por psicología popular se entiende el conjunto de presuposiciones que desde el sentido común tenemos de nuestras propias mentes y de las ajenas. Estas presuposiciones, que nos abstenemos de llamar ‘conocimiento’ para no dar por sentada la cuestión de su verdad, nos permiten explicar el comportamiento de los otros, entender lo que sucede en nuestro interior y, en breve, movernos en entornos sociales. Otra manera de definir a la psicología popular es la siguiente. Consiste en la atribución de estados

mentales como una herramienta para la explicación, predicción y control de la conducta (Churchland, 1991, Horgan & Woodward, 1985). Esta caracterización de la psicología popular se cuestiona al menos desde tres ángulos distintos.

Primero, desde quienes afirman que pensar a la psicología popular como una sola herramienta que se usa para varios fines es un error. Esta perspectiva surge de la investigación en psicología social que ha mostrado, entre otras cosas, que las habilidades de explicación y predicción de la conducta son asimétricas (Andrews, 2012, p. 37) y, aún más importante, que se pueden hacer consideraciones sobre la conducta de los otros más allá de la atribución de estados mentales, y esto se puede hacer a partir de aplicación de estereotipos, rasgos del carácter, comportamientos pasados, etc. (Andrews, 2012, p. 65).

Por otra parte, algunas investigaciones han mostrado que la atribución de estados mentales no intencionales (sensaciones, emociones, estados de ánimo) parece hacerse no con el estricto fin de hacer consideraciones sobre la conducta de los individuos, sino que se harían para tener un punto de partida para realizar ciertos juicios morales. Es decir, la atribución de sensaciones a un organismo, más que ser útil para predecir o explicar su comportamiento, sería útil como un insumo para la posterior realización de un juicio moral (Knobe & Prinz, 2008).

Finalmente, desde la filosofía experimental, se ha generado un debate respecto a la naturaleza de ciertos conceptos de la psicología popular que inicialmente parecían descriptivos, pero que, según algunos de sus usos, parecen tener un componente normativo. Un punto de partida aceptado de la discusión es que cuando usamos términos como "creencia", "deseo", "felicidad" o "enamoramiento", entre muchos otros, intentamos describir lo que sucede en las cabezas de las personas. Es decir, que describen un estado de cosas del mundo. Sin embargo, recientes exploraciones empíricas sobre el uso de varios de estos conceptos parecen mostrar que hay algo de normativo en ellos. Es decir, que cuando usamos conceptos como "felicidad" y "enamoramiento" (Phillips, et al., 2011) e "intencionalidad" (Knobe, 2003) no sólo describimos lo que sucede en el sujeto a quienes atribuimos el estado, sino que además estamos expresando cierto estándar normativo relativo a consideraciones más allá de las mentales, por ejemplo, qué tan valioso es el motivo por el que atribuimos felicidad, qué tan bueno o malo es un efecto de una acción o qué tanto merece alguien ser el objeto de amor de otro. Nuestro artículo se encargará de revisar la literatura experimental que justo se enmarca dentro de este último proyecto.

§2. EL EFECTO KNOBE

Joshua Knobe (2003) realizó el siguiente experimento. Le entregó a sus participantes una de dos historias. La primera historia, que llamaremos la condición de daño, es la siguiente:

El vicepresidente de una compañía fue donde el presidente de la junta y dijo: “estamos pensando en comenzar un nuevo programa. Nos ayudará a aumentar las ganancias, pero también dañará el medio ambiente”.

El presidente de la junta le respondió: “No me importa en absoluto dañar el medio ambiente. Sólo quiero tener tantas ganancias como podamos. Comencemos el nuevo programa”.

Ellos iniciaron el programa. Efectivamente, el medio ambiente fue perjudicado.

La segunda historia, la condición de ayuda, es similar a la anterior en todo, excepto que en ella el nuevo programa del que hablan los personajes no tiene como efecto colateral el daño al medio ambiente sino que por el contrario lo beneficia.

El vicepresidente de una compañía fue donde el presidente de la junta y dijo: “estamos pensando en comenzar un nuevo programa. Nos ayudará a aumentar las ganancias, pero también ayudará al medio ambiente”.

El presidente de la junta le respondió: “No me importa en absoluto ayudar el medio ambiente. Sólo quiero tener tantas ganancias como podamos. Comencemos el nuevo programa”.

Ellos iniciaron el programa. Efectivamente, el medio ambiente se benefició.

Con base en estas viñetas, los participantes debieron contestar a las preguntas: ¿el presidente dañó intencionalmente el medio ambiente?, en el primer caso o, ¿el presidente ayudó intencionalmente el medio ambiente?, en el segundo. Los sujetos de la condición de daño respondieron en su gran mayoría afirmativamente: el 77 por ciento de los participantes consideraron que el presidente sí había dañado intencionalmente el medio ambiente. En contraste, la gran mayoría de los sujetos que recibieron la viñeta de ayuda contestaron negativamente: el 70 por ciento consideró que el presidente no había ayudado intencionalmente al medio ambiente.

Esta asimetría es claramente anómala. Si la atribución de estados mentales, como la intención, se hace con el ánimo de describir lo que sucede, o ha sucedido, en la cabeza de alguien en virtud de su conducta observable, es extraño que en las dos condiciones donde el personaje ha tenido una idéntica conducta (ha dicho lo mismo, ha ordenado lo mismo), los participantes califiquen de manera opuesta la intencionalidad del efecto colateral. A esta asimetría se le conoce como el “efecto Knobe”.

La explicación de Knobe a esta anomalía ha resultado polémica. El resultado del efecto colateral en la condición de daño puede calificarse

como malo, mientras que en la condición de ayuda puede calificarse como bueno. Así, habría una diferencia en el carácter moral de ambos efectos. Según Knobe, ahí radica la explicación de la asimetría. El concepto de acción intencional se aplica a efectos colaterales cuando el resultado es moralmente reprochable, pero no se aplica cuando el resultado es positivo. De esta manera, el concepto no es completamente descriptivo, es decir, no sólo se usa para describir un estado mental, sino que también tiene un componente normativo. Se hace una evaluación del resultado del efecto respecto a un estándar moral y luego se aplica el concepto.

Lo polémico de la solución propuesta por Knobe radica en que ella estaría en contra de la visión estándar de la psicología popular. En particular en dos de sus rasgos: i) no sería un concepto enteramente descriptivo y, además, ii) no se aplicaría exclusivamente con base en las consideraciones de las conductas de los agentes a quienes se les atribuyen dichos estados.

§3. OTRAS EXPLICACIONES

Al abanico de explicaciones respecto al efecto Knobe podemos ubicarlas en un mapa donde se encuentran dos vectores. Por un lado, el vector horizontal, donde se diferencian las posiciones dependiendo de si las consideraciones morales son tomadas en cuenta o no. Como veremos, las explicaciones que desechan el aspecto moral como el aspecto relevante van a resultar más explicativas porque dan cuenta de un rango más amplio de evidencia, i.e., asimetrías en casos donde no hay ningún rasgo moral. El vector vertical donde distinguimos si la explicación implica decir algo acerca de la naturaleza del concepto o no dice nada acerca de ella. Justo esta distinción permitirá ubicar a quienes consideran que este tipo de evidencia dice algo acerca de la naturaleza del concepto de acción intencional y quienes consideran que no lo hace. El cuadro está en la página siguiente.

§3.1. CONSIDERACIONES MORALES QUE NO ESTÁN RELACIONADAS CON LA NATURALEZA DEL CONCEPTO (CUADRANTE SUPERIOR DERECHO)

Algunos críticos de Knobe concedieron que ciertas consideraciones morales explicaban la asimetría en las atribuciones del concepto en los casos de efectos colaterales. Sin embargo, agregaron que tales consideraciones morales no hacían parte del núcleo del concepto, es decir, que la evidencia del efecto Knobe no decía nada en absoluto sobre la naturaleza del concepto de acción intencional. Tales consideraciones morales son externas al concepto y de algún modo guían o desvían la aplicación de él. Ahora bien, la discusión es, entonces, a qué nivel ocurren esas consideraciones morales. Mele (2001) defiende que hay creencias morales explícitas que afectan

la atribución; Nadelhoffer (2004) considera que quienes cumplen este papel son ciertas emociones relacionadas con la atribución de culpa; Adams & Steadman (2004a, 2004b) proponen que las consideraciones morales ocurren a un nivel pragmático y, por último, Wright y Bengson (2009) explican la asimetría en virtud de otra: una asimetría en las consideraciones de elogio/culpa y la conexión que generalmente desde el sentido común tendría esto con la intencionalidad. A continuación examinaremos con más detalle estas cuatro hipótesis.



CUADRO 1.
Distribución de las propuestas según la clasificación que proponemos.

Antes de la difusión de los experimentos de Knobe, Mele (Mele & Sverdlik, 1996) se había mostrado en desacuerdo con la idea de que la atribución de intencionalidad estuviera comúnmente influenciada por consideraciones morales. Son abundantes los casos en donde al atribuir intencionalidad y valor moral los sujetos demuestran, a través de sus respuestas, que tácitamente entienden y atribuyen ambas cosas de manera independiente (Mele & Sverdlik, 1996, p. 269). Tres ejemplos de ello son:

i) un verdugo que acaba con la vida de un condenado no es considerado por la mayoría de la gente como alguien que haya hecho algo moralmente reprochable pero sí como alguien que lo hizo intencionalmente; ii) un conductor ebrio que arrolla y mata a un transeúnte, se considera generalmente como alguien que hizo algo malo aunque sin intención, y iii) alguien que por escuchar música a alto volumen no se percata que su hijo tuvo un accidente, allí se hace un reproche moral aunque no se le atribuye el haber tenido la intención de hacerlo. Así, pareciera que en principio ambas atribuciones se hacen de manera independiente.

Aunque, *prima facie*, estos casos muestran que las atribuciones de intencionalidad y moralidad son independientes, Mele sugiere que en la asimetría del efecto Knobe sí hay una relación entre ambas. Sería usual que desde el sentido común se tuvieran ciertos prejuicios que vinculan a estas dos atribuciones, y dichos prejuicios podrían variar de caso a caso y no son necesariamente consistentes entre sí. Además, tales prejuicios serían corregibles, es decir, a la luz de cierto examen y evaluación podrían cambiarse. Ejemplo de un prejuicio con estas características podría enunciarse de la siguiente manera: “para que alguien sea culpable por haber hecho algo, debió hacerlo intencionalmente”. Ahora bien, dado que los experimentos de Knobe se realizaron recolectando opiniones inmediatas de personas que comúnmente no reflexionan a fondo sobre estos temas, sus juicios sin depurar no necesariamente reflejan la naturaleza de los conceptos, sino de sus prejuicios. En resumen, la explicación de Mele (2001) al efecto Knobe consiste en afirmar que la asimetría en las respuestas se presenta porque los sujetos experimentales tienen ciertos prejuicios irreflexivos del sentido común que afectan la aplicación del concepto.

Por otra parte, Nadelhoffer (2004) sostiene que la asimetría sucede porque los sujetos que se enfrentan a las viñetas *sienten* que el agente es merecedor de culpa y tal sentimiento afecta su atribución de intencionalidad. En ambos escenarios el presidente de la compañía muestra desinterés por el medio ambiente y tal actitud es reprochable. De tal modo, en la condición de daño los sujetos tienden a atribuirle intencionalidad al efecto colateral, mientras en la condición de ayuda no lo hacen.

Un presupuesto empírico importante en la hipótesis de Nadelhoffer es el orden en que los juicios relacionados ocurren. Según la idea estándar (Knobe 2008, p. 138), un sujeto primero evalúa si la acción es buena o mala, luego si el agente lo hizo de manera intencional y en virtud de ambos se genera el juicio respecto a si quien está implicado es merecedor de culpa o no. Según lo propuesto por Nadelhoffer, sucede de otro modo. El sujeto experimental, al leer la viñeta, hace un juicio sobre la acción y también sobre la culpa que le corresponde al personaje. De ese modo, tal emoción que le lleva a considerar al personaje merecedor de culpa es la que afecta su juicio de atribución de intencionalidad.

Al continuar este cuadrante de respuestas tenemos la propuesta de Adams & Steadman (2004a, 2004b) que apela a consideraciones de corte pragmático. En la base de su propuesta, Adams & Steadman señalan que a nivel pragmático los sujetos ordinarios encuentran una equivalencia entre la expresión “*a* es culpable de haber hecho *x*” con “*a* hizo *x* a propósito” y “*a* hizo *x* intencionalmente”. Lo anterior indicaría que en el momento en que se le pregunta a los sujetos experimentales si creen que el presidente perjudicó el medio ambiente intencionalmente, ellos asocian la pregunta con: “¿considera que el presidente es culpable de dañar al medio ambiente?”. Lo cual afecta el patrón de respuestas. A partir de este hecho Adams & Steadman defenderán la tesis de que la asimetría en las atribuciones de intencionalidad por parte de sujetos experimentales se explica por algunas implicaturas pragmáticas.

Por implicaturas pragmáticas se entienden aquí todo aquello que se comunica y está sujeto al carácter contextual de una oración y no a la deducción que parte de su carácter semántico. Paul Grice muestra que hay conclusiones que se dan naturalmente al interior de una conversación que no se siguen lógicamente (Grice, 2011, p. 515). Por ejemplo, Susana dice: “hoy no deseo ver a Oscar”. Alguien podría fácilmente interpretar que Susana desea no ver a Oscar el día de hoy, aunque esto no se sigue. Para que esta persona interpretara correctamente lo que quiso decir, Susana debió haber dado más información. De igual manera, cuando alguien dice: “Susana hizo aquello a propósito”, usualmente se concluye “Susana es culpable de haber realizado aquello”, a pesar de que esto no se siga lógicamente.

De acuerdo con Adams & Steadman, los hablantes ordinarios no tienen un concepto articulado de acción intencional. Esta propuesta encuentra apoyo en estudios realizados por Knobe y Mele (Mele, 2001. Knobe y Nichols, 2003) cuyos resultados muestran que los participantes no coinciden en los criterios aparentemente básicos de la acción intencional, i.e., creencia, deseo, habilidad, intención o conciencia. No obstante, a pesar de que no se puede estudiar el núcleo del concepto de acción intencional, sí se pueden estudiar las características pragmáticas de tal concepto.

Cuando el participante se enfrenta a la viñeta y lee que el presidente de la compañía dice: “no me importa en absoluto dañar al medio ambiente” o “no me importa en absoluto ayudar al medio ambiente”, en ambos casos se genera un reproche. Según Adams & Steadman, lo que los participantes reprochan no es la previsión de que se va a dañar el medio ambiente que el presidente tiene, dado que el vicepresidente le ha informado al respecto. Lo que los participantes reprochan es la indiferencia del presidente hacia el medio ambiente. El presidente decide ignorar el daño.

En la condición de ayuda, la previsión es positiva, el vicepresidente informa al presidente de que su acción va a tener como efecto colateral el

beneficio al medio ambiente. Sin embargo, el presidente expresa su desinterés al respecto. De este modo, dado que él declara no estar interesado en ayudar al medio ambiente, el sujeto experimental no lo responsabiliza positivamente por el beneficio ocasionado. De tal modo, cuando se le pregunta si el presidente ayudó intencionalmente al medio ambiente, el participante responde negativamente.

Por último, la hipótesis de Wright & Bengson (2009) es que la asimetría que se percibe en el efecto Knobe se produce debido a dos factores: i) una asimetría en la atribución de elogio y culpa, y ii) el hecho de que la intencionalidad generalmente vincula la consideración moral de la acción con la responsabilidad de los actores. Veamos en detalle los dos componentes. La asimetría que Wright & Bengson encuentran en la atribución de elogio y culpa es la siguiente. Mientras que para la atribución de culpa (responsabilidad negativa) respecto a un resultado de una acción quien atribuye no toma en cuenta si es un efecto directo o un efecto colateral, para la atribución de elogio sí lo toma en cuenta. Para la atribución de elogio (responsabilidad positiva) es relevante que el resultado elogiable de la acción sea el efecto directo y no un simple efecto colateral. Es decir, el resultado a elogiar debe darse por las razones correctas, mientras que un resultado reprochable puede darse de cualquier modo. Esta asimetría es clara en el caso de las viñetas de “daño” y “beneficio”. En el caso de “daño” es intuitivo señalar como responsable negativamente al vicepresidente de la compañía, mientras que en el caso de “beneficio” no parece tan intuitivo señalarlo como digno de elogio.

El segundo componente de la hipótesis es el hecho de que generalmente se considera que hay una estrecha conexión entre intencionalidad y responsabilidad. “Normalmente, aquellos que actúan intencionalmente para obtener un resultado malo son negativamente responsables y aquellos quienes actúan intencionalmente para obtener un resultado bueno son positivamente responsables” (Wright & Bengson, 2009, p. 27; la traducción es nuestra). Esto no quiere decir que la intencionalidad sea necesaria, por ejemplo, el caso del conductor ebrio, donde la negligencia reemplaza la intencionalidad.

Se pueden expresar estas relaciones así:

Acción buena/mala + acción intencional = actor positivamente/negativamente responsable.

Normalmente, a partir de la evaluación del resultado en adición con la presencia o ausencia de intencionalidad inferimos la presencia o ausencia de responsabilidad positiva o negativa en el agente. Sin embargo, como en toda buena ecuación, también se podría inferir la presencia o ausencia de intencionalidad si tenemos la evaluación del resultado de la acción y también la atribución de responsabilidad. Justo esto es lo que explicaría el

efecto Knobe. El sujeto experimental se enfrenta a la historia, en la viñeta de “daño”, evalúa el resultado de la acción como negativo y además, al saber que el personaje conocía el resultado de implementar el programa, le atribuye responsabilidad negativa por el efecto colateral. Una vez que tiene esos dos elementos, infiere que el personaje realizó el efecto colateral de manera intencional.

§3.2. CONSIDERACIONES MORALES QUE ESTÁN
RELACIONADAS CON LA NATURALEZA DEL CONCEPTO
(CUADRANTE INFERIOR DERECHO)

La explicación de Knobe se ubica en este cuadrante. Como se presentó en §2, Knobe considera que su evidencia muestra algo importante acerca del núcleo del concepto. De algún modo el concepto de acción intencional no es un concepto enteramente descriptivo sino que también tiene un componente normativo. Es decir, la asimetría en la aplicación del concepto se debe a que éste no sólo describe la situación sino que incluye un estándar normativo (en este caso moral) que determina su aplicación.

§3.3. CONSIDERACIONES NO MORALES QUE NO ESTÁN
RELACIONADAS CON LA NATURALEZA DEL CONCEPTO
(CUADRANTE SUPERIOR IZQUIERDO)

De acuerdo con Edouard Machery (2008) la discusión que estamos revisando presupone la distinción propuesta por Chomsky (1965) entre *desempeño* y *competencia* sobre un determinado concepto. La competencia es el conocimiento que tiene el hablante del concepto. Mientras que el desempeño es el uso real del concepto en situaciones determinadas. Es de resaltar que tal uso puede ser afectado por condiciones externas al concepto: limitaciones cognitivas (i.e., memoria, percepción), contextos conversacionales, estados emocionales, entre otros.

Ahora bien, la posición de Knobe es que la asimetría muestra algo acerca de la naturaleza del concepto, mientras que las propuestas de Mele, Nadelhoffer y Adams & Steadman justamente señalan que hay un error en el nivel del desempeño. Machery quiere demarcarse de esos términos de la discusión. Dado que la distinción entre esos dos niveles no ha sido claramente demarcada, Machery sostiene que no se puede decir mucho respecto a si esta asimetría dice algo o no acerca de la naturaleza del concepto. Más bien, se centra en una explicación de la asimetría en términos de mecanismos psicológicos subyacentes. Explicación que no implica nada acerca de la naturaleza del concepto y, por tanto, está ubicado en este cuadrante.

Según Machery (2008), la asimetría se explica porque los sujetos califican como intencional todo efecto colateral que puede ser considerado como un costo que se asume con el fin de obtener un beneficio. El presidente de la compañía asumió el costo de dañar el medio ambiente con el fin de

obtener más ganancias. Caso contrario sucede en la condición de “ayuda”, beneficiar al medio ambiente no puede ser entendido como un costo.

La definición de “costo” implica dos condiciones: i) es todo aquello que se requiere para alcanzar un fin, y ii) se valora negativamente (Machery, 2008, p. 176). Si bien en ambas viñetas el efecto colateral puede satisfacer la condición (i), en la condición de beneficio la viñeta no satisface (ii), por lo tanto, no es calificado como intencional. Según la hipótesis del balance [*trade-off hypothesis*] de Machery, la atribución de intencionalidad va de la mano con una evaluación de costos.

Un aspecto interesante de la teoría de Machery es que saca la explicación de la asimetría del núcleo del concepto. Machery no sugiere que para ser competente con el concepto de acción intencional se deba ser competente evaluando costos. Más bien, que la discusión acerca del efecto Knobe no dirá nada acerca de la naturaleza del concepto, y que en el mejor de los casos llegará a decirnos algo acerca del uso del concepto en un momento determinado. La evaluación de costos de la que habla facilita la comprensión del patrón de aplicación del concepto. No dice nada acerca del mismo.

Es importante en esta teoría que en tanto no apela a ninguna consideración moral puede explicar otras asimetrías en la atribución de intencionalidad que no parecen tener ningún componente moral. Tal es el caso de la anomalía que ocurre en los escenarios del vaso promocional gratuito. Esta asimetría parece no poder explicarse con base en consideraciones morales, dado que la lectura de la viñeta no suscita ninguna. Estas son las viñetas:

José se sentía bastante deshidratado, así que paró en una tienda para comprar un refresco en la presentación más grande posible. Antes de ordenar, el cajero le dijo que el refresco extragrande ahora costaba un dólar más de lo que costaba antes. José le contestó: “no me importa si tengo que pagar un dólar más, yo sólo quiero la bebida más grande que tengas”. Efectivamente, José recibió el refresco extragrande y pagó un dólar más por él.

José se sentía bastante deshidratado, así que paró en una tienda para comprar un refresco en la presentación más grande posible. Antes de ordenar, el cajero le dijo que si compraba un refresco extragrande, se lo darían en un vaso conmemorativo especial. José le contestó: “no me importa lo del vaso conmemorativo, yo sólo quiero la bebida más grande que tengas”. Efectivamente, José recibió el refresco extragrande en el vaso conmemorativo.

A los participantes se les preguntó si José pagó intencionalmente el dólar adicional y el 95 por ciento opinaron que sí. Mientras que cuando se les preguntó si José recibió intencionalmente el vaso conmemorativo, sólo el 45 por ciento respondió afirmativamente. Aquí hay una diferencia en la

atribución de intencionalidad que no puede ser respondida por Knobe o por aquellas otras hipótesis que apelan a consideraciones morales.

En contraste, la hipótesis de Machery sí puede dar cuenta de ello. En el escenario del dólar extra, los sujetos experimentales atribuyen intencionalidad porque se incurre en un costo para recibir un beneficio. Por el contrario, en el caso del vaso conmemorativo no hay ningún costo en el que se incurra para tener el beneficio de la bebida, y así los participantes tienden a no considerar esta acción como intencional.

Mencionaremos dos problemas de la hipótesis de Machery. Primero, Machery deliberadamente no deja en claro si es necesario que el personaje de la viñeta entienda el costo como tal o si es posible que el personaje no lo asuma. Es decir, es suficiente con que el sujeto experimental considere que un daño al medio ambiente es un costo, o se necesita que el sujeto experimental considere que el presidente de la compañía toma el daño al ambiente como un costo. Esta diferencia que a primera vista puede ser sutil en realidad es importante. Machery en la nota al pie 10 se refiere al asunto como algo secundario. Sin embargo, esto podría hacer la diferencia. Pareciera que en el caso del presidente de la compañía, su declaración: "No me importa en absoluto dañar el medio ambiente", puede tomarse como una expresión de indiferencia que indica que para él eso no es un costo. Así que el punto acerca de quién debe verlo como un costo no es irrelevante.

Un segundo problema, fue señalado por Mallon (2008). Realizó un experimento con las siguientes viñetas:

Un miembro de una pandilla local fue donde el líder y dijo: "Estamos pensando en intentar una nueva táctica. Con ella inundaremos el barrio de cocaína barata, incrementando nuestras ganancias, pero también perjudicaremos a los policías dado que más policías morirán a causa de la violencia relacionada.

El líder respondió: "Admito que estaría bien perjudicar a los policías, pero realmente no me interesa. Lo que quiero es tener tantas ganancias como sea posible. Vamos a implementar la nueva táctica".

Ellos implementaron la nueva táctica, y efectivamente, los policías se perjudicaron dado que murieron más de ellos en la violencia relacionada con drogas.

La segunda viñeta fue:

Un miembro de una pandilla local fue donde el líder y dijo: "Estamos pensando en intentar una nueva táctica. Con ella inundaremos el barrio de cocaína barata, incrementando nuestras ganancias, pero también ayudaremos a los más adictos a tener dinero para su comida y renta.

El líder respondió: "Admito que estaría bien para estas personas tener más dinero para comida y renta, pero realmente no me interesa. Lo que quiero es tener tantas ganancias como sea posible. Vamos a implementar la nueva táctica".

Ellos implementaron la nueva táctica, y efectivamente, los más adictos se beneficiaron al tener más dinero.

Un 62 por ciento de los participantes, afirmaron que el líder de la pandilla perjudicó intencionalmente a los policías, mientras que sólo un 28 por ciento afirmó que el líder de la pandilla ayudó intencionalmente a los adictos. Este caso, según Mallon, presenta una objeción a la hipótesis del balance. En las viñetas se encuentra una asimetría que no puede ser explicada con base en una evaluación de costos dado que ninguno de los dos efectos, perjudicar a los policías o ayudar a los adictos, puede ser entendido como un costo. De tal modo, según Mallon, la hipótesis de Machery no explica satisfactoriamente todas las asimetrías en la atribución del concepto de acción intencional.

A pesar de estas dificultades, consideramos que la hipótesis de Machery aún puede mantenerse en pie. Frente a la primera, aunque Machery quiera permanecer apartado de la cuestión de quién debe ser el evaluador del costo, consideramos que se podría tomar una decisión al respecto sea vía análisis de las viñetas o sea haciendo un estudio experimental. A primera vista, parece que el costo debiera ser analizado en términos del sujeto experimental y no del personaje de la viñeta. Cuando el presidente de la compañía afirma: “no me importa en absoluto dañar el medio ambiente” declara que para él eso no es un costo. Sin embargo, el participante podría pensar que aunque el presidente no lo considere así, de hecho, dañar el medio ambiente sí es un costo que el presidente asume, pues éste se verá afectado de algún modo. Sobre si esta interpretación funciona con las demás viñetas habría que hacer un análisis más extenso que no haremos aquí. Un estudio más detallado incluiría solicitar a sujetos experimentales hacer una evaluación de costos explícita.

Respecto a la prueba realizada por Mallon, la asimetría en las viñetas sobre el líder de la pandilla sí pueden ser explicadas desde la hipótesis del balance. Mientras que ayudar a los más adictos no genera nada negativo para el líder de la pandilla, asesinar más policías sí podría hacerlo. A diferencia de Mallon, consideramos que el participante podría hacer una evaluación más amplia pensando, por ejemplo, que asesinar más policías traerá más represión policial, más enfrentamientos, más investigaciones, etc., situaciones que a fin de cuentas tendrán un impacto negativo en las actividades de la pandilla. Sobre este punto tendría que hacerse trabajo empírico adicional.

§3.4. CONSIDERACIONES NO MORALES QUE ESTÁN RELACIONADAS CON LA NATURALEZA DEL CONCEPTO (CUADRANTE INFERIOR IZQUIERDO)

Dentro de este cuadrante ubicamos las hipótesis de Bence Nanay (2010) y de Chandra Sripada (2010). Ambos proponen hipótesis que terminan por

enriquecer el concepto de acción intencional de una manera que no hubiera sido previsible sino gracias a la investigación empírica. Sin embargo, en ninguno de los dos casos proponen que hayan cuestiones de moralidad dentro del concepto. Para Nanay, es más un asunto de evaluación del estatus modal de las acciones. Mientras que para Sripada hay una evaluación de concordancia entre los rasgos psicológicos estables del personaje, el Yo Profundo [*Deep Self*], y el comportamiento del personaje, el Yo Activo.

La tesis de Nanay es que el efecto Knobe se explica no porque al aplicar el concepto de acción intencional los sujetos experimentales tomen en cuenta consideraciones morales, sino consideraciones modales. En resumen, la asimetría en la aplicación del concepto se debe a que en la condición de “daño” el sujeto establece, de manera no consciente, que si el personaje no hubiera tenido las mismas consideraciones (es decir, si fuese cierto que le hubiera importado perjudicar el medio ambiente) el resultado de la acción sería distinto. Sin embargo, en la condición de ayuda si el personaje no hubiera tomado las mismas consideraciones (es decir, si fuese cierto que le hubiera importado beneficiar el medio ambiente) entonces el resultado de la acción sería el mismo. Es en este aspecto modal, sobre lo que podría haber pasado, donde se encuentra la explicación. Cuando el resultado pudo haber sido distinto si le hubiera importado el efecto colateral, entonces éste se califica como intencional. Pero si es el caso que a pesar de importarle el efecto colateral la acción arrojaría el mismo resultado, entonces se tiende a pensar que tal efecto colateral no fue intencional. En palabras de Nanay: “Se juzga un efecto colateral previsto como intencional si la siguiente afirmación modal es correcta: si ella no hubiera ignorado sus consideraciones acerca del efecto colateral previsto, su acción pudo haber sido diferente (en igualdad de condiciones)” (Nanay, 2010).

Nanay defiende que su tesis es superior ya que explica la evidencia de Knobe, de Machery y la de Mallon. Además, explica los resultados de su propia prueba, que es esta:

El vicepresidente de una compañía fue donde el presidente de la junta y le dijo: “estamos pensando en expandir nuestro edificio. Esto nos ayudará a incrementar nuestras ganancias, pero también significa que tenemos que cortar el árbol de manzanas que está al frente de su oficina”.

El jefe de la junta responde: “aunque yo pasé mucho tiempo trepando el árbol como un niño y disfrutando de su vista desde la oficina, no me importa en absoluto cortarlo. Sólo quiero tener tantas ganancias como podamos. Comiencen la ampliación.”

Ellos comenzaron la expansión. El árbol de manzanas fue cortado.

En la otra condición, el presidente de la compañía declara que el árbol le molesta desde que se cambió a esa oficina. Dice que eso igual no le importa, ordena ampliar el edificio y el árbol de manzanas fue cortado. En el primer caso la mayoría (78 por ciento) dice que cortó el árbol intencionalmente. En el segundo la mayoría (69 por ciento) dice que no lo cortó intencionalmente.

Según Nanay, no se podría explicar esto desde la teoría de Knobe porque el resultado de ambos casos es el mismo: se corta el árbol, así que tendría que tener la misma valencia moral. La diferencia entre ambas condiciones, que explicaría la asimetría, es respecto a las razones que podría haber tenido el presidente de la compañía para actuar de otro modo. En el primer escenario, el presidente tenía una razón para no cortar el árbol. En el segundo escenario no la tenía. De tal modo, la teoría de Nanay sí explica este resultado. En el primer caso, si el presidente de la compañía no hubiera ignorado sus consideraciones respecto al efecto secundario sus acciones habrían sido distintas, habría atendido sus buenos recuerdos con el árbol y no hubiera ordenado la ampliación de oficinas. En el segundo caso, si el presidente no hubiera ignorado sus consideraciones respecto al efecto secundario, su acción sería la misma, dado que tenía razones para cortar el árbol.

La hipótesis de Sripada parte de que los participantes concluyen ciertos rasgos estables de carácter del personaje en cuestión. A estos rasgos les llama el Yo Profundo [*Deep Self*]. Sripada menciona como rasgos del Yo Profundo a los valores, actitudes, principios, personalidad y cualquier otro rasgo psicológico estable que no varíe a pesar de las situaciones. Ahora bien, la asimetría ocurre porque entra en juego una evaluación de los sujetos respecto a si hay una concordancia entre la acción realizada por el personaje y el Yo Profundo del mismo.

Cuando los sujetos experimentales leen en la viñeta que el presidente de la compañía expresa: “no me importa en absoluto el medio ambiente”, puede deducir ciertos rasgos de su carácter, por ejemplo, que sólo le importa el dinero, es egoísta y desconsiderado. Si esto es así, entonces aquello que estaría en concordancia con el Yo Profundo del presidente sería no beneficiar el medio ambiente y por lo tanto en la condición de “daño” tienden a responder que dañar el medio ambiente fue una acción intencional. En la condición de “ayuda”, el presidente se mantiene en la misma situación de indiferencia respecto al medio ambiente, pues su situación respecto a su carácter es la misma. Si esto es así, entonces al beneficiar el medio ambiente no hay concordancia entre su carácter y sus acciones, *ergo*, no es una acción intencional.

§4. CONCLUSIONES

En esta revisión hemos presentado ocho hipótesis que buscan dar cuenta de las asimetrías en la aplicación del concepto de acción intencional en los casos presentados. Como conclusión presentaremos un breve balance acerca de qué soluciones nos parecen más poderosas que otras.

Un primer aspecto que nos parece simple, pero tal vez no ha sido suficientemente explícito en la discusión, es que las propuestas que apelan a consideraciones morales tienen un poder explicativo menor que aquellas que no lo hacen. ¿La razón? Hay pruebas que han mostrado asimetrías en escenarios donde es difícil hacer una lectura moral. Por ejemplo, ¿qué podría decir la hipótesis de Nadelhoffer (2004) respecto al escenario del vaso gratis? Tal hipótesis señala que un sentimiento de atribución de culpa es el origen de la anomalía en la atribución. Sin embargo, en esa viñeta no habría nadie a quién culpar de nada. Lo mismo sucede con las hipótesis de Knobe, Adams & Steadman y la de Wright & Bengson. Dado que la historia no tiene ningún rasgo moral relevante, tales hipótesis resultan inertes explicativamente.

Esto nos deja con que las hipótesis de mayor capacidad explicativa son justamente aquellas que no hacen referencia a ningún rasgo moral de las situaciones. Es decir, aquellas que están al lado izquierdo en el mapa propuesto en el apartado §3. Así tenemos que son las hipótesis de Machery, Sripada y Nanay aquellas que en principio parecen resultar mejores opciones. Frente a estas tres, la hipótesis de Sripada no logra explicar unos resultados que las otras dos hipótesis sí logran explicar.

En el caso del árbol de manzanas presentado por Nanay, al presidente de una compañía se le pide autorización para cortar un árbol que está en el centro del edificio para poder realizar una ampliación de las instalaciones que traerá como consecuencia un aumento en las ganancias. En los dos escenarios, el presidente de la compañía declara no importarle el árbol y ordena que se corte. Sin embargo, en uno de los casos el presidente declara que ese árbol le gustaba y pasaba tiempo en él, mientras que en el otro el presidente afirma que siempre le molestó el árbol. Los resultados que presenta Nanay (2010) es que hay una asimetría cuando se les pregunta a los participantes si consideran que el presidente de la compañía cortó el árbol intencionalmente. En el primer caso tienden a afirmar que lo hizo intencionalmente, y en el segundo sucede lo contrario.

Frente a este caso, la hipótesis de Sripada (2010) fracasa. Dado que en ambas condiciones el presidente de la compañía expresa lo mismo, que el árbol de manzanas no le importa, es de esperar que el Yo Profundo que los participantes le atribuyen sea el mismo en ambos casos: un hombre que sólo le preocupa el dinero, que no le importa el medio ambiente, que incluso puede pasar por encima de su bienestar por alcanzar más ganancias. En ambos casos, el resultado de las acciones también es el mismo. El

presidente autoriza la ampliación del edificio y por consiguiente se corta el árbol. Dado que las acciones son las mismas y los rasgos a partir de los cuales se infiere el Yo Profundo son los mismos, y estos dos están en concordancia, la predicción que puede hacerse desde la hipótesis del Yo Profundo de Sripada es que en ambos casos los sujetos experimentales dirán que el presidente cortó el árbol intencionalmente. Sin embargo, justo lo que sucede es lo contrario. Sólo en uno de los casos los sujetos respondieron así, y en el otro lo hicieron del modo opuesto. Así, la hipótesis de Sripada no puede dar cuenta de lo sucedido.

Quedamos entonces con la hipótesis de Nanay y de Machery. Ya vimos en la sección §3.4. cómo Nanay (2010) explica esa evidencia. Consideramos que Machery (2008) también brinda elementos para poder explicar la asimetría. En el primer escenario, el presidente de la compañía asume algo negativo (atentar contra algo que lo hacía feliz) en aras de realizar el objetivo primario de la acción (realizar la ampliación); de este modo, podemos decir que el presidente de la compañía asume un costo. Por tal razón, en ese escenario la mayoría de participantes califica el efecto colateral como producto intencional. En el segundo escenario, el presidente de la compañía no asume ningún costo, dado que a él en realidad no le gustaba el árbol. Por tal motivo, en ese escenario no se califica como intencional el efecto secundario. De cualquier manera, tendremos que hacer trabajo experimental adicional para corroborar si efectivamente las personas ven en el primer escenario un costo. Hasta el momento sólo señalamos que tal interpretación es perfectamente plausible.

En síntesis, el criterio que proponemos para ubicar las tesis de Nanay (2010) y Machery (2008) como las más promisorias para ser las explicaciones definitivas del efecto Knobe, es que ellas explican toda la evidencia que ha sido presentada. Nueva evidencia aparece con frecuencia, así que vale la pena estar atentos. Por otra parte, queda abierta la cuestión de si tal criterio es correcto. Hay voces, que bien podrían ser calificadas como pluralistas o pesimistas, que afirman que ninguna teoría podrá dar cuenta de toda la evidencia (Nichols & Ulatowsky, 2007) dado que el término "intencional" es polisémico y nunca sabremos con seguridad cuál de los significados es el que los participantes en cierto momento están tomando en cuenta. Si bien tal posición pluralista tiene sus virtudes, en especial evitar la frustración de no encontrar una teoría que explique todos los datos, creemos que aún podemos esperar hallar una explicación que abarque la evidencia en su totalidad.

BIBLIOGRAFÍA

- Adams, F. & Steadman, A. (2004a), "Intentional action in ordinary language: core concept or pragmatic understanding?" *Analysis* 64: 173-181.
- Adams, F. & Steadman, A. (2004b), "Intentional action and moral considerations: still pragmatic," *Analysis* 64: 268-276.
- Andrews, K. (2012), *Do Apes Read Minds? Toward a New Folk Psychology*, Cambridge: MIT Press.
- Chomsky, N. (1965), *Aspects of a Theory of Syntax*, Cambridge: MIT Press.
- Fajardo- Chica, D. (2011), "Incendiando el sillón", en *El Hombre y la máquina*, Universidad Autónoma de Occidente 36: 16-27.
- Knobe, J. (2003), "Intentional actions and side effects in ordinary language," *Analysis* 63: 309-324.
- Knobe, J. & Prinz, J. (2008), "Intuitions about consciousness: experimental studies", *Phenomenology and the Cognitive Sciences* 7 (1): 67-83.
- Knobe, J. (2008), "The concept of intentional action: a case study in the uses of folk psychology," in Knobe, J. & Nichols, S. (eds.) *Experimental Philosophy*, Oxford: OUP.
- Horgan, T. & Woodward, J. (1985), "Folk psychology is here to stay," *The Philosophical Review* 94: 197-226.
- Machery, E. (2008), "The folk concept of intentional action: philosophical and experimental issues," *Mind and Language* 23: 165-189.
- Mallon, R. (2008), "Knobe vs. Machery: testing the trade-off hypothesis," *Mind and Language* 23: 247-255.
- Mele, A. & Sverdlik, S. (1996), "Intention, intentional action, and moral responsibility," *Philosophical Studies* 82(3): 265-287.
- Mele, A. (2001), "Acting intentionally: probing folk notions," in Malle, B. F., Moses, I. J. & Baldwin, D. (eds.), *Intentions and Intentionality: Foundations of Social Cognition*, Cambridge: MIT Press.
- Nadelhoffer, T. (2004), "Praise, side effects, and intentional action," *Journal of Theoretical and Philosophical Psychology* 24: 196-213.
- Nadelhoffer, T. & Nahmias, E. (2007), "The past and future of experimental philosophy," *Philosophical Explorations* 10(2): 123-149.
- Nanay, B. (2010), "Morality or modality: What does the attribution of intentionality depend on?" *Canadian Journal of Philosophy* 40(1): 25-39.
- Nichols, S. (2004), "Folk concepts and intuitions: from philosophy to cognitive science," *Trends in Cognitive Science* 8(11): 514-518.
- Phillips, J., Misenheimer, L. & Knobe, J. (2011), "The ordinary concept of happiness (and others like it)," *Emotion Review* 71: 929-937.
- Sripada, C. (2010), "The deep self model and asymmetries in folk judgements about intentional action," *Philosophical Studies* 151(2): 159-176.
- Wright, J. & Bengson, J. (2009), "Asymmetries in judgements of responsibility and intentional action," *Mind and Language* 24(1): 24-50.