



Los parámetros normativos de corrección racional implícitos en las investigaciones en filosofía experimental

El caso del experimento que correlaciona la receptividad al bullshit con un déficit cognitivo

María Natalia Zavadvivker*

Introducción

En este artículo me propongo introducir una serie de reflexiones en torno del peso normativo de los supuestos axiológicos y epistémicos implícitos en las investigaciones en filosofía experimental. Para tal fin, he tomado como ejemplo paradigmático el experimento llevado a cabo por Pennycook, Cheyne, Barr, Koehler y Fugelsang, publicado en un *paper* de 2015 bajo el título “On the reception and deception of pseudoprofound bullshit”. Brevemente, el objetivo del experimento de Pennycook et al. (2015) era hallar correlaciones entre la tendencia a aceptar o rechazar frases “bullshit”, y la posesión de ciertas habilidades cognitivas asociadas a dos estilos de razonamiento diferenciales: uno más intuitivo, consistente en una excesiva apertura mental y una alta predisposición a dotar de significatividad, profundidad y trascendencia a cualquier declaración; y otro más analítico, consistente en la adopción de una posición escéptica inicial que favorecería la detección de falacias e inconsistencias en el discurso como resultado de su sometimiento a un escrutinio que involucraría procesos deliberativos lentos. El término “bullshit” es un neologismo traducible como “basura pseudoprofunda”, y alude a aquellos mensajes que circulan asiduamente en redes sociales, consistentes en frases grandilocuentes y presuntamente profundas con palabras de moda asociadas a la autoayuda que pretenden ser estimulantes, trascendentes y motivadoras pero que en rigor carecen de significado y valor de verdad. Los autores pretendían analizar los factores que intervienen tanto en la receptividad como en la sensibilidad al bullshit, entendiendo por ‘receptividad’ la aceptación y valoración positiva de tales frases, y por ‘sensibilidad’ la disposición opuesta, vale decir, la tendencia a someterlas a un escrutinio crítico que permita

* Universidad Nacional de Tucumán (UNT), CONICET.
zavadvivker@gmail.com

detectar su carácter falaz. Para tal fin solicitaron a los participantes que evalúen frases pseudoprofundas en una escala de profundidad, con el propósito de establecer una medida legítima de receptividad al “bullshit” (BSR). El objetivo era correlacionar esa clasificación con diferencias individuales en la posesión de dos estilos diferenciales de pensamiento: uno más analítico (asociado a una actitud de escepticismo inicial que favorecería el reclutamiento de habilidades cognitivas asociadas a la detección de conflictos), y otro más intuitivo, que según los autores podría estar asociado, por un lado, a una fuerte tendencia a aceptar a-priori cualquier información recibida como verdadera o significativa (excesiva apertura mental que predispondría a encontrar profundidad y significatividad en cualquier discurso) y, por el otro, a un déficit cognitivo conocido como “fallas en el monitoreo de conflictos”, que contribuiría a que el receptor confunda vaguedad o falta de claridad con profundidad. También conjeturaron que las personas más receptivas al bullshit serían más propensas a las confusiones ontológicas (errores categoriales usuales en el pensamiento *folk* en los que se confunden los planos animado e inanimado, físico y mental, etc. –por ejemplo, cuando se comparten cadenas de oración para velar por la salud de alguien, asumiendo que las oraciones tienen capacidades curativas–), y a las creencias epistémicamente sospechosas (aquellas que entran en conflicto con concepciones naturalistas comunes sobre el funcionamiento del mundo, como las creencias religiosas y paranormales, las teorías conspiracionistas y la aceptación de terapias medicinales alternativas). Los autores esperaban hallar correlaciones entre este tipo de creencias y la receptividad al bullshit.

En el primer experimento se pidió a 280 estudiantes universitarios que evalúen en una escala del 1 al 5 (1: para nada profunda, 5: muy profunda) diez declaraciones sintácticamente correctas, pero formadas por una selección aleatoria de términos vagos en boga. Las frases fueron tomadas de dos sitios web: uno que mezcla al azar palabras comúnmente utilizadas en los tweets de Deepak Chopra, y otro que genera frases compilando términos que “suenan profundos”. El supuesto era que las calificaciones altas indicarían alta receptividad al bullshit. Los participantes completaron además cinco tareas tendientes a evaluar diferencias individuales en el estilo cognitivo analítico y en las habilidades cognitivas: la prueba de reflexión cognitiva (problemas matemáticos que inducen respuestas intuitivas incorrectas); una batería de “heurísticos y sesgos”, en donde también se inducen respuestas intuitivas incorrectas basadas en heurísticas o prejuicios

comunes; y dos medidas de habilidad cognitiva: la prueba Wordsum sobre inteligencia verbal; y un test sobre habilidades aritméticas globales. Para evaluar las confusiones ontológicas, los participantes debían responder cuán literal o metafórico les parecía el contenido de una serie de frases. Quienes las percibieran como más literales que metafóricas presuntamente estarían evidenciando una mayor confusión ontológica. Finalmente, completaron un cuestionario sobre su nivel de acuerdo/desacuerdo con 8 creencias religiosas comunes (más allá, cielo, infierno, milagros, ángeles, demonio, alma, Satanás).

En el segundo estudio repitieron el experimento utilizando ejemplos reales de bullshit (“tweets” particularmente vagos de la cuenta de Twitter de Deepak Chopra). También incluyeron medidas de auto-reporte sobre la disposición al pensamiento analítico e intuitivo, e incorporaron un test extra de inteligencia fluida.

En el tercer estudio compararon la escala de profundidad atribuida a las frases “bullshit” con la misma escala atribuida a citas genuinamente motivacionales e inteligibles, conjeturando que las diferencias en la atribución de profundidad a frases legítimamente significativas y a frases pseudoprofundas les permitiría obtener una medida fiable de sensibilidad al bullshit. También pidieron a los sujetos que evaluaran frases mundanas con un significado claro pero que no consideraríamos profundo (por ej.: “Mucha gente disfruta con tal tipo de música”).

En el cuarto estudio se correlacionó la BSR con una escala de creencias paranormales, otra de creencias conspiracionistas, y otra sobre el grado de aceptación de diez tipos de medicinas alternativas. También implementaron una escala de personalidad que mide las diferencias individuales en cinco rasgos (extroversión, amabilidad, responsabilidad, estabilidad emocional y apertura).

Entre los resultados más relevantes, los autores hallaron correlaciones significativas entre la propensión a adjudicar profundidad a las declaraciones “bullshit” y la mayoría de las pruebas que medían habilidades cognitivas asociadas al pensamiento analítico: los más receptivos fueron menos reflexivos, evidenciaron una menor inteligencia fluida, verbal y aritmética, y demostraron mayor propensión a las confusiones ontológicas, las ideas conspiracionistas, las creencias religiosas y paranormales y la medicina alternativa. Para medir específicamente la sensibilidad al bullshit (capacidad de detectar ambigüedad, inconsistencias y conflictos cognitivos), los investigadores compararon las evaluaciones que los par-

ticipantes hicieron de frases bullshit y de citas motivacionales legítimamente significativas. Encontraron correlaciones entre dicha sensibilidad, el estilo cognitivo analítico y el escepticismo en relación a los fenómenos paranormales, pero no pudieron correlacionarla con las ideas conspiracionistas ni con la aceptación de tratamientos medicinales alternativos. Sí encontraron correlaciones robustas entre la receptividad al bullshit y las creencias epistémicamente dudosas, y detectaron la propensión de un buen número de personas a una mente excesivamente abierta asociada a un débil espíritu crítico (incluso algunos participantes adjudicaron cierta profundidad a frases completamente mundanas). También hallaron una correlación negativa entre la predisposición a creer en entidades sobrenaturales (creencias religiosas) y dos factores complementarios: una apertura mental excesiva que insta al receptor a adjudicar profundidad y valor espiritual a cualquier estímulo; y la inhibición de la capacidad analítica y reflexiva, ya que el sujeto no quiere “ver” las contradicciones implicadas en las creencias religiosas y las explicaciones paranormales.

Mientras en mi ponencia anterior he sugerido la necesidad de incorporar a este experimento un conjunto adicional de variables de orden pragmático que en situaciones de la vida real podrían influir en una mayor o menor receptividad al bullshit (por ej., la velocidad requerida para la decodificación de las frases, la reputación -intelectual y moral- del emisor, el efecto de sugestión producido por la aceptación, aprobación y viralización de la frase en las redes sociales, el estado anímico del receptor de la frase, su nivel sociocultural, etc.); en esta ponencia me propongo:

- Realizar algunas observaciones críticas a los supuestos teóricos y axiológicos implícitos en el diseño experimental de Pennycook y colaboradores.
- Utilizar al presente estudio como un ejemplo de los escollos que deben enfrentar los filósofos experimentales cuando procuran obtener evidencia objetiva sobre el pensamiento de sentido común (intuiciones filosóficas *folk*, entendiendo por tales las creencias explícitas de la gente y los procesos cognitivos subyacentes), pero bajo la asunción de posiciones axiológicas y normativas implícitas en relación a cuáles son las creencias epistémicamente “correctas” y los mecanismos cognitivos adecuados para obtenerlas.

Algunas observaciones al diseño experimental

En su trabajo, los autores definen aproximadamente el “bullshit” como frases consistentes en conjuntos aleatorios de palabras motivacionales de moda (las cuales suelen despertar reacciones emocionales asociadas a la profundidad y la trascendencia) con una estructura sintáctica correcta, pero que carecen de sentido y, por ende, en principio serían difíciles de decodificar. A su vez, procuran correlacionar la receptividad al bullshit con la propensión a incurrir en confusiones ontológicas y sustentar creencias epistémicamente dudosas. Ahora bien, si bien es posible que personas propensas a la religiosidad, a creer en fenómenos paranormales, inclinarse por terapias alternativas o teorías conspiracionistas, etc., estén al mismo tiempo más abiertas a adjudicar cierta profundidad a frases ininteligibles; no es lo mismo una declaración perfectamente inteligible cuyo contenido denote una confusión ontológica (Por ej.: “La fuerza de la oración salvará al mundo”), o creencias epistémicamente dudosas (“Todos tenemos un ángel de la guarda que nos protege”), que una frase carente de sentido por haber sido elaborada apelando a una conjunción aleatoria de términos (Por ej.: “Si llegamos a maximizar nuestras experiencias transparentes estaremos preparados para monetizar las nuevas sinergias globales” –extraída de un generador de bullshit en español–). Por lo general en la vida real las personas tienden más a compartir, postear, viralizar y darle “like” a frases motivacionales claramente inteligibles pero cuyas afirmaciones desafían las leyes naturales y el conocimiento de sentido común, que a frases difícilmente decodificables que en última instancia no terminan sirviendo al fin supuestamente perseguido: generar estados motivacionales positivos y provocar efectos de sanación espiritual apelando a las emociones del receptor (cabe señalar además que su objetivo no es proporcionar información objetiva y con valor de verdad sobre cierto estado de cosas). Probablemente los autores, para justificar su presunción de que mucha gente es muy receptiva a frases pseudoprofundas ininteligibles, supusieron que lo que puede estimularlas es el impacto emocional que desencadenan algunos términos motivacionales de moda, los cuales podrían por sí solos disparar sentimientos de profundidad y trascendencia, aun cuando la frase completa carezca de sentido. Sin embargo, a mi juicio cuanto más clara e inteligible es una frase, mejor cumple su fin de fomentar la autoayuda y la psicología positiva. En otras palabras,

Pennycook et al. (2015) asociaron implícitamente la ininteligibilidad y oscuridad de las frases con el valor de verdad de su contenido, determinado por la inclusión de confusiones ontológicas y creencias epistémicamente dudosas. Esto sugiere que los autores adhieren implícitamente a ciertas ramas de la filosofía analítica ligadas al Positivismo Lógico (por ejemplo, la filosofía del primer Wittgenstein) que identifican el significado de un término o proposición con sus condiciones de verdad (vale decir, con sus posibilidades, directas o indirectas, de contrastación empírica). De allí que correlacionen la apertura y aceptación positiva de frases “carentes de sentido”, con la propensión a las confusiones ontológicas y a las creencias epistémicamente dudosas. No queda muy claro si los autores sugieren que estas frases carecen de sentido porque no sería lícito o legítimo suponer la posibilidad de que, por ej., los planos natural y sobrenatural, o material y espiritual se superpongan en los hechos (con lo cual estarían asociando sentido o inteligibilidad a valor de verdad), o si simplemente quieren indagar si las personas propensas a las creencias sobrenaturales (por tener una apertura mental excesiva, y, por ende, una menor predisposición al pensamiento analítico) también son más propensas a aceptar el bullshit. Sea como sea, al procurar ofrecer una definición aproximada de “bullshit”, los autores parecen asociar “significado” a “valor de verdad” cuando afirman que tales frases carecen de ambas cosas, siendo que dicha asociación es privativa de cierto posicionamiento filosófico. Por ej., esta posición sería compatible con la filosofía analítica del primer Wittgenstein, mientras que dicho autor en su segunda etapa asume un criterio mucho más amplio, según el cual el significado de los términos está determinado por su uso en ciertos contextos y por la atribución intersubjetiva de sentido que le adjudica un grupo determinado, siempre en relación a ciertas formas de vida. Mi propósito no es cuestionar la posición de los investigadores, sino poner de manifiesto en qué medida los filósofos experimentales inevitablemente deben partir de supuestos metafísicos, axiológicos y normativos que les permitan establecer de antemano ciertos parámetros de “corrección” en virtud de los cuales determinan, de manera inconsciente e implícita, qué tipo de intuiciones y procesos cognitivos recabados experimentalmente serán correctos y cuáles serán “irracionales”, por alejarse de sus propios estándares. En otras palabras, las investigaciones de la filosofía experimental no se agotan, como es su pretensión originaria, en la mera descripción neutral de las intuiciones filosóficas *folk* de la gente, sino que dicha descripción presupone como telón de fondo juicios norma-

tivos y de valor implícitos que involucran criterios acerca de lo correcto e incorrecto (en este caso en términos cognitivos), lo que es incompatible con esa especie de retorno al empirismo ingenuo que los filósofos experimentales sustentan con bastante convicción. El supuesto de base (que aunque a juicio personal me resulte sensato, es como mínimo discutible) es que existe un ínfimo sector de la población –el de los filósofos o intelectuales formados en el pensamiento crítico– que estaría en condiciones de determinar qué tipo de creencias y procesos cognitivos serían correctas o lícitas, y cuáles serían erróneas y fundadas en razonamientos falaces. En el fondo no se trata de una asunción tan descabellada, pues del mismo modo podríamos decir que son los conceptos teóricos de los físicos, y no las intuiciones físicas *folk* de la gente común, los que explican adecuadamente las leyes y variables del mundo físico. Incluso para Wittgenstein (1921) la tarea específica de la filosofía es precisamente la de esclarecer y depurar los conceptos del lenguaje ordinario, bajo el supuesto de que este proceso nos habilitaría para razonar correctamente acerca del mundo. Pero estos supuestos deberían ser al menos enunciados explícitamente en las investigaciones en filosofía experimental, dejando en claro la posición metafísica, epistémica, axiológica e ideológica de los autores, de modo tal que éstos asuman su punto de partida subjetivo y circunscripto a una perspectiva determinada, en lugar de presentar supuestos datos crudos analizados desde una presunta posición neutral.

¿Es posible evitar asunciones normativas implícitas en la descripción de las intuiciones filosóficas *folk*?

Si bien los autores del experimento son psicólogos, su trabajo puede ser enmarcado dentro de la línea de investigación conocida como “filosofía experimental”. Los filósofos experimentales procuran recabar datos empíricos sobre el pensamiento espontáneo de las personas en relación a ciertos tópicos filosóficos (posibles intuiciones filosóficas *folk*, y los procesos cognitivos responsables de su fijación), pero para evaluar tales intuiciones (por ej., para ponderar su grado de corrección e incorrección), inevitablemente deben establecer ciertos parámetros sobre la forma correcta de razonar filosóficamente (procesos cognitivos fiables) y sobre la corrección de las respuestas emitidas, lo que implica que el abordaje descriptivo sólo puede llevarse a cabo asumiendo ciertos supuestos normativos fijados por los propios filósofos. Esto implica la suposición de que son los especialis-

tas en filosofía, y no la mayoría estadística de las personas, quienes estarían legítimamente autorizados para juzgar el valor de las intuiciones y procesos cognitivos del resto de los sujetos. Tal como lo mencionamos anteriormente, en algún sentido esto se justifica, ya que los filósofos son precisamente esa ínfima franja de la población dedicada al entrenamiento profesional y sistemático de habilidades cognitivas asociadas al pensamiento analítico, con lo cual poseen un estilo cognitivo más reflexivo que el promedio. Sin embargo, una cosa es tener una mayor disposición al pensamiento reflexivo, y por ende, un mayor entrenamiento en las habilidades cognitivas asociadas; y otra es suponer que dicho entrenamiento garantiza la adquisición de las creencias “correctas”, o al menos mucho más fiables que las obtenidas por otras vías (por ej., las emociones, dado que estas últimas tenderían a ser sesgadas y falaces). Si bien en lo personal adhiero a esta posición, y considero adecuada la asociación entre una excesiva apertura mental, escasa deliberación crítica y reflexiva; y el pensamiento “mágico” (creencias religiosas y paranormales, supersticiones, etc.); es precisamente el racionalismo crítico del filósofo el que nos insta a cuestionar la validez de nuestras propias creencias y a asumir las dificultades en la búsqueda de parámetros valorativos universales y objetivos para juzgarlas (de hecho, nada garantiza la falsedad de las creencias religiosas y paranormales, ya que en sentido estricto son irrefutables al ser empíricamente incontrastables). Pero, al mismo tiempo, el filósofo experimental debe adoptar algún parámetro con el cual cotejar el grado de “corrección” de las creencias y procesos cognitivos de los sujetos experimentales, aun cuando el propósito de la investigación sea meramente descriptivo. En el experimento analizado la propia adopción del término “bullshit” (traducible al español aproximadamente como “pura mierda”, vocablo con una carga axiológica profundamente despectiva) sugiere un claro posicionamiento ideológico, de modo que los criterios para identificar frases “bullshit” no pueden agotarse en un conjunto de rasgos claramente demarcables y meramente denotativos capaces de evitar la subjetividad del evaluador. Esta incorporación de valoraciones subjetivas en el diseño experimental es un claro síntoma del “giro pragmático” de la filosofía (influido además por sus cruces interdisciplinarios con dominios tradicionalmente pertenecientes a otras ciencias): los filósofos van abandonando cada vez más sus recaudos profesionales ligados a la neutralidad axiológica y el análisis puramente lógico de los conceptos, para asumir abiertamente ciertas posiciones (de manera consciente o no) amparados en criterios pragmáticos

que los instan a no suspender excesivamente el juicio, a fin de no disociar la investigación metateórica de sus asunciones ideológicas y valorativas sustentadas de hecho. Si bien los autores del estudio son psicólogos cognitivos y no filósofos, lo que podría explicar su menor compromiso con la neutralidad axiológica y la imparcialidad en el uso de los términos, los filósofos experimentales también tienden a adoptar (más o menos implícitamente) posiciones axiológicas y normativas como parámetros para evaluar las evidencias obtenidas, aunque algunos asumen su problematización y las controversias epistemológicas implicadas.

En síntesis, ninguna evidencia empírica puede dar cuenta *per se* de cuán apropiadas o inapropiadas son las intuiciones *folk* si no se adopta de antemano un criterio normativo que permita valorarlas, y es el experimentador quien se arroga tal potestad. El análisis de los resultados consiste entonces en cotejar la evidencia obtenida con parámetros previamente establecidos sobre tipos de respuestas adecuadas o inadecuadas, correctas o erróneas. En el experimento analizado, por ej., lo correcto es sostener una posición escéptica y crítica en relación a ciertas declaraciones, no sólo por su vaguedad, ambigüedad y presunta ininteligibilidad, sino porque su contenido supone la aceptación de creencias que desafían las leyes naturales y el conocimiento avalado empíricamente. Pero la valoración de los datos empíricos depende también de ciertas asunciones teóricas más globales sobre el status de la cognición humana. Por ej., una posición pragmática de corte evolucionista supone que nuestras creencias y herramientas cognitivas son adaptaciones instrumentalmente útiles para lidiar con presiones ambientales específicas, incluyendo desafíos relativos a la necesidad de lograr un equilibrio emocional y espiritual, lo que abonará a favor de la búsqueda de explicaciones y la fijación de creencias capaces de dotar de sentidos y significados trascendentes a nuestra existencia. Desde esta perspectiva, la tendencia a aceptar confusiones ontológicas y creencias epistémicamente dudosas podría considerarse “apropiada” por resultar eficaz para el logro de algún fin adaptativo. Pero si asumimos desde una perspectiva racionalista que las intuiciones y procesos cognitivos deben cumplir con ciertos estándares de racionalidad independientemente de su valor adaptativo o utilidad práctica, entonces el dictamen de la mayoría no influirá en absoluto en el veredicto a priori respecto de qué intuiciones o procesos cognitivos son más adecuados que otros.

Referencias

- Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. (2015). On the reception and deception of pseudoprofound bullshit. *Judgment and Decision Making*, *10*(6), 549–563.
- Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72.
- Wittgenstein, L. (1921). Logisch-philosophische Abhandlung. *Annalen der Naturphilosophie*, *14*, 185–262.