

Razonamiento deductivo en entornos colaborativos: un análisis de la evaluación de argumentos sobre el aborto

Massolo, Alba^{*, a, b}  & Traversi, Mariel^b 

Artículo Original   

Resumen

Este estudio tiene como objetivo analizar la resolución individual y la resolución colaborativa de una tarea deductiva con contenido ideológicamente controvertido. Para esto, diseñamos y realizamos una prueba de evaluación de argumentos. Todos los argumentos utilizados tenían conclusiones a favor o en contra del aborto. Los resultados obtenidos muestran que, en sintonía con nuestra primera hipótesis, el desempeño colaborativo en la resolución de esta tarea deductiva es superior al desempeño individual. Sin embargo, se observa una excepción en la resolución de los silogismos inválidos increíbles. Asimismo, respecto de los modelos de combinación social, los datos reflejan que, a diferencia de lo que sucede con la resolución de otras tareas deductivas y de nuestra segunda hipótesis, el modelo la verdad gana no resulta ser el modelo que mejor describe la resolución colaborativa de esta tarea deductiva. Discutimos estos hallazgos a partir de las teorías sobre el origen social del razonamiento humano.

Palabras clave:

psicología del razonamiento, razonamiento deductivo, razonamiento colaborativo, sesgos cognitivos, sesgo a mi favor

Abstract

Deductive reasoning in collaborative settings: An analysis of argument evaluation on abortion This study aims to analyse individual and collaborative resolution of a deductive task with ideologically controversial content. To this end, we designed and conducted an argument evaluation test. All the arguments analyzed had conclusions in favour of or against abortion. The results show that, in line with our first hypothesis, collaborative performance in solving this deductive task is superior to individual performance. However, an exception is observed in the resolution of invalid unbelievable syllogisms. Additionally, regarding social combination models, the data indicate that, unlike the resolution of other deductive tasks and our second hypothesis, the truth-wins model is not the one that best describes the collaborative resolution of this deductive task. We discuss these findings in light of theories about the social origins of human reasoning.

Keywords:

psychology of reasoning, deductive reasoning collaborative reasoning, cognitive biases, myside bias.

Tabla de Contenido

Introducción	46
Método	49
Resultados	50
Discusión	54
Referencias	56

Recibido el 27 de mayo de 2025; Aceptado el 14 de octubre de 2025

Editaron este artículo: Mariano Rodríguez, Débora Burin, Andrea Garofolo, Anabella Dichiará, Sebastián Miranda y Maylin Martínez Muñoz

El razonamiento deductivo permite establecer una conexión de carácter necesario entre las premisas y la conclusión de un argumento. Así, si las premisas de un argumento deductivamente válido son verdaderas, resulta imposible que su conclusión sea falsa. Tradicionalmente, en psicología del razonamiento se ha adoptado la lógica clásica como estándar normativo para evaluar el razonamiento deductivo humano (Over & Evans, 2024). Siguiendo este paradigma de análisis, la psicología del razonamiento ha demostrado a lo largo de los años que el desempeño humano en tareas de razonamiento

deductivo es bastante pobre, es decir, que los seres humanos encuentran grandes dificultades para resolver tareas deductivas simples (Evans, 2002). En este sentido, abundan los estudios empíricos que muestran los errores del razonamiento humano en la resolución de diversos problemas deductivos (Evans et al., 1983; Stanovich & West, 2008; Thompson et al., 2003; Wason, 1968). Sin embargo, los resultados de este tipo de estudios cambian notablemente cuando estas mismas tareas deductivas son resueltas de manera grupal y colaborativa (Trouche et al., 2014). Estos datos sugieren que el

^a Universidad Nacional de Córdoba, Facultad de Filosofía y Humanidades, Córdoba, Argentina.

^b Universidad Católica de Córdoba, Facultad de Ciencias de la Salud, Córdoba, Argentina.

*Enviar correspondencia a: Massolo, A. E-mail: alba.massolo@unc.edu.ar

Citar este artículo como: Massolo, A., & Traversi, M. (2026). Razonamiento deductivo en entornos colaborativos: un análisis de la evaluación de argumentos sobre el aborto. *Revista Argentina de Ciencias del Comportamiento*, 18(1), 46-57

razonamiento deductivo en entornos colaborativos es superior al razonamiento individual.

Estos resultados empíricos están en sintonía con las teorías filosóficas que abogan a favor del origen dialógico de la deducción. En este sentido, se ha argumentado que el hecho de que los seres humanos sean mejores razonadores deductivos cuando tienen la oportunidad de intercambiar razones y argumentos en un entorno dialógico constituye un fuerte punto a favor de la concepción dialógica de la deducción (Dutilh-Novaes, 2021). Asimismo, las teorías evolucionistas sobre los orígenes sociales del razonamiento humano ofrecen una justificación filogenética a la emergencia de estas habilidades deductivas. Desde estas teorías, se sostiene que los humanos comenzaron a desarrollar habilidades cognitivas de agencia colectiva cuando las características de su medioambiente los llevaron a colaborar con otros miembros de su especie a fin de tener éxito en la búsqueda de alimentos (Tomasello, 2022). Esto dio lugar a la coordinación con otros individuos, distinguiendo roles y perspectivas, para alcanzar los objetivos compartidos. En una línea de investigación afín, se ha propuesto que, partiendo de un kit genético inicial, los humanos desarrollaron en sus prácticas culturales una serie de dispositivos cognitivos que se transmitieron culturalmente (Heyes, 2018). A partir de estos dispositivos cognitivos, habrían emergido las prácticas argumentativas que a la vez dieron el sustento para el desarrollo del razonamiento deductivo (Dutilh-Novaes, 2021).

Una de las tareas más emblemáticas para analizar el razonamiento deductivo humano es el test de Wason (Wason, 1968). En esta tarea deductiva, que consiste en presentar a las y los participantes cuatro cartas y pedirles que decidan qué carta o cartas deberían seleccionar a fin de determinar la verdad de una afirmación condicional, se intenta analizar si las personas razonan siguiendo las reglas de inferencia del condicional material. Es decir, lo que se espera a partir del estándar marcado por la lógica es que las personas usen una estrategia de confirmación siguiendo el *modus ponens* y una estrategia de falsación siguiendo el *modus tollens*. Sorprendentemente, los resultados muestran que apenas un 9% de las personas seleccionan las cartas correctamente (Johnson-Laird & Wason, 1970). A partir de este resultado, el test de Wason comenzó a replicarse en una gran cantidad de

estudios, siempre obteniendo un porcentaje de respuestas correctas por debajo del 10% (Evans, 2005). Sin embargo, en una de las réplicas de este test, se propuso que la resolución de esta tarea deductiva se realizara también de manera grupal. Los resultados mostraron un significativo incremento en el porcentaje de respuestas correctas de un 75% en la resolución grupal, frente a un 9% de respuestas correctas en la resolución individual (Moshman & Geil, 1998). Posteriormente, los siguientes estudios realizados sobre la resolución grupal de esta tarea de razonamiento mostraron siempre un mejor desempeño deductivo grupal en comparación con el desempeño individual (Boku et al., 2018; Mercier et al., 2016). Asimismo, la resolución grupal colaborativa de otras tareas de razonamiento deductivo, como la tarea de Levesque (1986), mostraron también una significativa mejora en relación con la resolución individual (Castelain et al., 2016; Mercier & Claidière, 2022; Trouche et al., 2014). De esta manera, estos estudios permiten concluir que el intercambio argumentativo dialógico, donde las personas participan junto con sus pares en el juego de dar y pedir razones, compartiendo argumentos, contraargumentos, dudas, críticas y justificaciones (Dutilh-Novaes, 2021; Köymen et al., 2020), incrementa significativamente el desempeño deductivo humano.

En un conjunto de investigaciones vinculadas a la resolución grupal de problemas, se ha planteado una serie de modelos de combinación social para dar cuenta de los procesos de toma de decisiones involucrados en la resolución de tareas intelectivas, es decir, tareas que tienen una única solución demostrablemente correcta dentro de un sistema matemático, lógico, científico o conceptual (Laughlin, 2011). Claramente, las tareas de razonamiento deductivo son equiparables a las tareas intelectivas puesto que tienen una única respuesta correcta dentro del sistema lógico correspondiente. De esta manera, siguiendo este planteo, es posible distinguir cuatro modelos de combinación social para los procesos de decisión colectiva. En primer lugar, el modelo de mayoría indica que una mayoría de respuestas individuales correctas, ya sea una mayoría simple o de dos tercios, es necesaria y suficiente para ofrecer una respuesta grupal correcta. En segundo lugar, el modelo de equiprobabilidad sugiere que la

respuesta correcta de todo el grupo tiene la misma probabilidad entre todas las respuestas dadas por los subgrupos compuestos por al menos un integrante. En tercer lugar, el modelo denominado la verdad respaldada gana indica que para que el grupo ofrezca una respuesta correcta son necesarias y suficientes dos respuestas correctas individuales de miembros del grupo. Y, finalmente, el modelo denominado la verdad gana sostiene que una única respuesta correcta dada por un miembro del grupo es necesaria y suficiente para que todo el grupo resuelva el problema de manera correcta (Laughlin & Ellis, 1986). En este marco, se ha mostrado que el modelo que mejor describe la decisión colectiva en tareas deductivas es el modelo la verdad gana. Sin embargo, también se ha reportado que incluso en grupos en los cuales ninguno de sus integrantes obtuvo la respuesta correcta de manera individual, en la instancia de razonamiento colectivo, trabajando de manera conjunta, estos grupos logran resolver el problema correctamente (Trouche et al., 2014).

En un estudio previo (Massolo et al., 2024), se analizó la resolución grupal de una tarea de evaluación de argumentos deductivos. En esta tarea, se les propuso a las y los participantes evaluar una serie de argumentos deductivos, tanto válidos como inválidos, con conclusiones creíbles e increíbles. Se trata de la típica tarea deductiva empleada para analizar el sesgo de creencia (belief bias), a saber, la tendencia sistemática a considerar válidos a los argumentos con conclusiones creíbles e inválidos a los argumentos con conclusiones increíbles, independientemente de su forma lógica (Béghin & Markovits, 2024; Boissin et al., 2023; de Chantal et al., 2020; Evans et al., 1983). La validez lógica es una propiedad formal de los argumentos según la cual, si las premisas son verdaderas, la conclusión necesariamente también lo es. Dado que la validez depende de la estructura y no del contenido, la verdad fáctica de las premisas es independiente de la validez del argumento. Así, un argumento puede ser válido, aunque su conclusión sea fáctica o empíricamente falsa, o ser inválido a pesar de tener una conclusión verdadera.

En sintonía con los estudios previos, en esta investigación se mostró que el desempeño grupal en la resolución de la tarea deductiva fue superior al desempeño individual, en tanto la resolución grupal obtuvo un mayor porcentaje de respuestas

correctas. Asimismo, los resultados obtenidos fueron consistentes con el modelo de combinación social denominado la verdad gana (Massolo et al., 2024). En el mencionado estudio, se utilizaron argumentos deductivos con contenido fáctico acerca del mundo, es decir, contenido que, al menos en principio, puede ser verificado o refutado por medios empíricos (Stanovich, 2021).

Así como el sesgo de creencia se vincula con las propias creencias sobre los hechos fácticos del mundo, existe otro fenómeno cognitivo similar, denominado sesgo a mi favor (myside bias), vinculado con las creencias derivadas de la propia ideología (Vedejová & Čavojeová, 2021). El sesgo a mi favor da cuenta de la tendencia sistemática a producir y aceptar argumentos que están en consonancia con las posturas ideológicas propias y a rechazar u objetar argumentos que contradicen la propia ideología. De esta manera, mientras el sesgo de creencia es un fenómeno que se observa en el razonamiento con contenido fáctico sobre el mundo, el sesgo a mi favor está vinculado con el contenido ideológicamente controvertido. Un rasgo notable del sesgo a mi favor, a diferencia de lo que sucede con otros sesgos cognitivos, es que no puede predecirse a partir de las medidas estándar de habilidades cognitivas tales como la inteligencia, las funciones ejecutivas y las disposiciones de pensamiento (Stanovich & Toplak, 2023). Estas características hacen del sesgo a mi favor un fenómeno cognitivo bastante atípico, con importantes consecuencias políticas y sociales (Stanovich, 2021).

El objetivo del presente estudio es analizar la resolución individual y la resolución grupal colaborativa de una tarea de evaluación de argumentos deductivos con contenido ideológicamente controvertido. Para esto, se diseñó y llevó a cabo una prueba de evaluación de argumentos que estuvo compuesta por un total de ocho silogismos categóricos. Estos ocho argumentos tenían conclusiones a favor y en contra del aborto. Como la prueba de razonamiento fue tomada a una muestra de estudiantes de Argentina, y en este país recientemente se dio un fuerte debate en torno al aborto, que culminó con su legalización a finales del año 2020, esta temática resulta óptima para un estudio de este tipo. A partir de los resultados obtenidos, se propone analizar, por un lado, si el desempeño grupal en la resolución de esta tarea

de evaluación de argumentos es superior al desempeño individual. Y, por otro lado, si el modelo la verdad gana es el modelo de combinación social que mejor describe la resolución colaborativa de esta tarea deductiva. Así, las dos hipótesis que guían esta investigación son las siguientes:

Hipótesis 1: El desempeño colaborativo en la resolución de la tarea deductiva es superior al desempeño individual.

Hipótesis 2: El modelo la verdad gana se aplica en la resolución grupal de la tarea deductiva.

Método

Participantes

Se empleó un muestreo circunstancial por conveniencia, con la participación de estudiantes de la carrera de psicología de una universidad argentina. Para estimar el tamaño muestral adecuado se utilizó el software estadístico *G*Power*, resultando necesario una muestra de por lo menos 45 participantes ($f = .25$, $\alpha = .05$). La muestra total estuvo compuesta por 111 participantes, con edades comprendidas entre 18 y 47 años ($M = 19.65$, $DE = 3.45$). La mayoría se identificó con el género femenino (81.3%), mientras que un 18.7% se identificó con el género masculino. Las y los estudiantes participaron voluntariamente, sin recibir compensación alguna.

Diseño

Se utilizó un diseño intra-sujetos de medidas repetidas en el cual los participantes debían determinar la validez de una serie de silogismos en tres instancias diferentes. Se establecieron tres condiciones: (1) resolución individual, (2) resolución colaborativa y (3) resolución individual posterior a la colaborativa. Estas tres condiciones (1, 2 y 3) se trataron como niveles de un factor intra-sujetos. Se consideraron como variables dependientes la cantidad de respuestas correctas (acierto/error), el tipo de silogismo (VC: válido creíble; VI: válido increíble; IC: inválido creíble; II: inválido increíble) y la postura frente al aborto (pro-aborto o anti-aborto). Asimismo, se incluyeron también como variables dependientes los índices de sesgo a mi favor, de creencia y de lógica.

Materiales

Encuesta de opinión

Previamente a la tarea de evaluación de argumentos, se administró una encuesta de

opinión basada en una escala Likert de seis puntos, con el objetivo de conocer la postura de cada participante sobre el aborto en lo relativo a sus decisiones y puntos de vista personales. Se evaluó la consistencia interna de esta escala mediante el coeficiente alfa de Cronbach, obteniéndose un valor alto ($\alpha = .93$). La encuesta solicitaba al comienzo que se indicara la edad y el género autopercebido.

Tarea de razonamiento

La tarea de evaluación de argumentos estuvo compuesta por ocho silogismos con contenido ideológico, específicamente sobre aborto. Cuatro eran válidos y cuatro inválidos, entre ellos cuatro con conclusiones pro-aborto y cuatro con conclusiones anti-aborto. Las opciones de respuesta eran: 'Sí, se sigue lógicamente' o 'No, no se sigue lógicamente'.

Procedimiento

Primero, los participantes aceptaron el consentimiento informado. Luego, completaron datos sociodemográficos y respondieron la encuesta de opinión. Antes de iniciar la tarea de evaluación de silogismos, se les entregaron instrucciones generales sobre cómo resolver los problemas, junto con una breve explicación del concepto de validez lógica y ejemplos de argumentos válidos e inválidos. Se instruyó a los participantes para que asumieran temporalmente que las premisas eran verdaderas y juzgaran únicamente si la conclusión se derivaba lógicamente de esas premisas. Se enfatizó explícitamente que no debían tener en cuenta el contenido de las premisas para decidir y sólo razonaran sobre la base de la relación lógica entre las premisas y la conclusión.

El estudio se desarrolló en tres fases consecutivas. En la primera fase, los participantes realizaron la tarea de evaluación de argumentos de manera individual, eligiendo entre las dos opciones para determinar si la conclusión se seguía lógicamente o no. Esta etapa corresponde a la condición 1, que representa la resolución individual. En la segunda fase, los participantes fueron agrupados de forma azarosa en grupos de tres a cinco personas y se les asignó la tarea de resolver los problemas de manera colectiva. En un primer momento, se buscó conformar grupos de cinco integrantes, siguiendo el criterio propuesto por Moshman y Geil (1998). Sin embargo, esta disposición no pudo sostenerse debido a las

limitaciones en el número total de participantes y a que la tarea se administró en distintas instancias. En consecuencia, se optó por conformar grupos de un mínimo de tres y un máximo de cinco integrantes. De este modo, la muestra quedó organizada en 16 grupos de cinco integrantes, cuatro grupos de cuatro integrantes y cinco grupos de tres integrantes. Se pidió a los participantes que llegaran a un consenso sobre cada argumento. Esta fase corresponde a la condición 2. Finalmente, en la tercera fase, se solicitó nuevamente a los participantes que completaran la tarea deductiva de manera individual. Esta última fase corresponde a la condición 3. En todas las fases se utilizó la misma tarea de razonamiento deductivo.

Análisis de datos

Los análisis estadísticos descriptivos e inferenciales se realizaron con el *software SPSS* versión 25. En primer lugar, se obtuvieron estadísticos descriptivos a partir de los datos demográficos ofrecidos por los participantes. Luego, se calcularon frecuencias y porcentajes de aceptación (silogismos donde se elige la opción “La conclusión sí se sigue lógicamente de las premisas”) y de respuestas correctas para cada tipo de silogismo. A fin de analizar estadísticamente los efectos de lógica y creencia, así como la interacción entre ellas, se calcularon los índices de lógica, creencia e interacción tomados de Evans y Curtis-Holmes (2005). Se llevaron a cabo pruebas t para evaluar si cada uno de los índices calculados era significativamente diferente de cero. Lo mismo se realizó con los índices de sesgo a mi favor (Čavojová et al., 2018). A su vez, se realizaron ANOVAs de medidas repetidas con cada uno de los índices, para identificar diferencias de medias entre las tres condiciones del estudio. Por último, se calcularon ANOVAs de medidas repetidas para comparar la cantidad de respuestas correctas totales entre las condiciones y, luego, diferenciando por tipo de silogismo.

Aspectos éticos

Antes de comenzar el estudio, se pidió a las y los participantes que leyeran y aceptaran el consentimiento informado. Asimismo, se les proporcionó información de contacto de las coordinadoras de la investigación, a fin de que puedan realizar cualquier pregunta o consulta una vez finalizado el estudio. Se les garantizó a los

participantes la confidencialidad de los datos recopilados y se les informó la posibilidad de retirarse del estudio en cualquier momento, si así lo deseaban.

El presente estudio forma parte de los Proyectos de Investigación, “Aspectos normativos del razonamiento humano: un abordaje filosófico y cognitivo” y “Lógica y psicología del razonamiento: un análisis filosófico y empírico sobre sesgos cognitivos”, que cuentan con el aval del Comité de Ética de la Facultad de Filosofía y Humanidades de la Universidad Nacional de Córdoba y del Comité de Bioética de la Universidad Católica de Córdoba, respectivamente.

El estudio siguió los lineamientos éticos de la *American Psychological Association* (2020). Todos los integrantes de la muestra participaron de manera voluntaria. Se siguieron las pautas de la Declaración de Helsinki (*Asociación Médica Mundial, 2004*) y la normativa de protección de datos.

Resultados

En primer lugar, y para determinar la posición ideológica de cada participante respecto del aborto, se estableció un punto de corte en los puntajes obtenidos en la encuesta de opinión. Para establecer el punto de corte, en primer lugar, se invirtieron las puntuaciones correspondientes a los ítems de orientación anti-aborto. Esto implicó que, por ejemplo, si un participante asignaba una puntuación de 6 a una afirmación en contra del aborto, dicha puntuación se transformaba en 1. Esta recodificación permitió unificar la dirección de todas las respuestas, de manera que los puntajes más altos reflejaran una postura pro-aborto y los más bajos una postura anti-aborto. A continuación, se calculó el rango posible de puntuaciones, cuyo mínimo fue 6 y máximo 36. Al dividir dicho rango por dos se obtuvo el valor 21, que se adoptó como punto de corte. En consecuencia, los participantes con una puntuación total superior a 21 fueron clasificados como pro-aborto, mientras que aquellos con un puntaje inferior a 21 fueron considerados anti-aborto. Como resultado, el 9% de los participantes se clasificó con una postura anti-aborto y el 91% con una postura pro-aborto.

De esta manera, los silogismos fueron clasificados como creíbles o no creíbles según la postura ideológica manifestada por cada participante en la encuesta de opinión. Es decir, si

un participante tenía una postura pro-aborto y el silogismo presentaba una conclusión válida a favor del aborto, este se etiquetó como válido creíble (VC). En cambio, si el participante se oponía al aborto, el mismo silogismo se consideró válido increíble (VI). Del mismo modo, un silogismo inválido con una conclusión anti-aborto se reclasificó como inválido increíble (II) para aquellos que apoyaban el aborto, y como inválido creíble (IC) para los que se oponían al aborto. Una vez que los silogismos se reorganizaron conforme a esta nueva clasificación, calculamos los índices de lógica, creencia e interacción, basándonos en Evans (2017) de la siguiente manera:

Índice de lógica: $(VC + VI) - (IC + II)$

Índice de creencia: $(VC + IC) - (VI + II)$

Índice de interacción: $(VI + IC) - (VC + II)$

El puntaje de cada índice se obtiene a partir de la suma de las proporciones de aceptación de los ocho silogismos según su categoría, con un rango posible que va de -2 a +2. Los índices aportan información clave. En primer lugar, la diferencia entre los silogismos válidos e inválidos permite evaluar el impacto de la lógica; un índice más alto indica una mayor corrección lógica en las respuestas. En segundo lugar, la variación en la aceptación de los silogismos creíbles frente a los increíbles refleja el efecto de las creencias; cuanto mayor es el índice, más pronunciado es el sesgo de creencia. Finalmente, el índice de interacción entre lógica y creencia señala el grado en que el sesgo de creencia es más evidente en silogismos inválidos que en válidos, aumentando conforme lo hace el índice.

Se realizaron pruebas *t* para evaluar la significancia estadística de la desviación de cada índice respecto de cero. Se observaron resultados positivos para todos los índices, excepto para el índice de creencia en la condición 2 ($p = .543$), lo que indica un efecto menor y no significativo de la creencia en la condición colaborativa. Además, se llevaron a cabo dos análisis de varianza de medidas repetidas (ANOVAs) para comparar los índices de lógica y creencia entre las tres condiciones. Se observó un efecto principal de la lógica ($F_{(1, 91)} = 33.28, p < .001, \eta^2 = .236$). Las comparaciones múltiples post hoc utilizando medias marginales con corrección de Bonferroni revelaron diferencias significativas. Específicamente, el índice de lógica fue significativamente más alto en la condición 2 ($M = 1.06, DE = .14$), en comparación con la condición

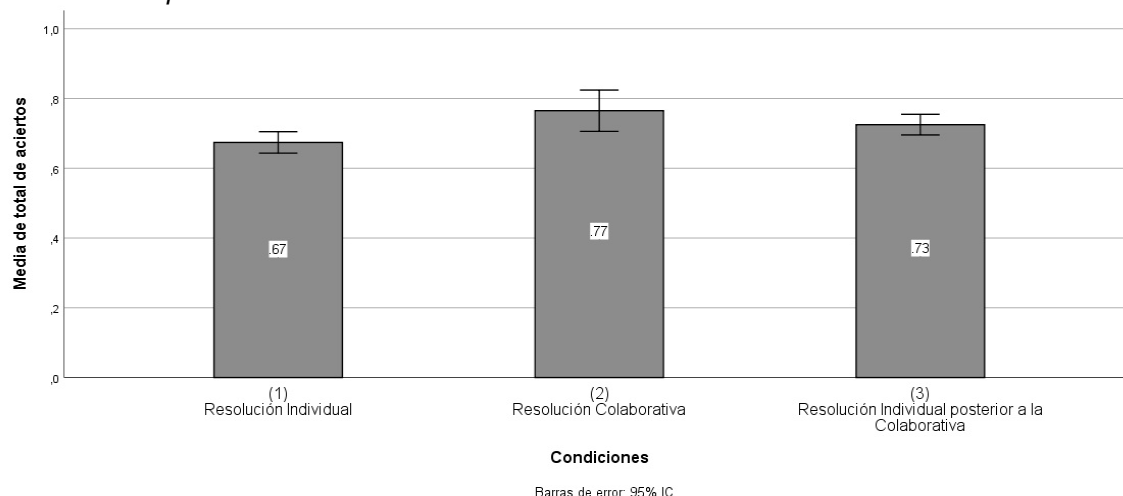
1 ($M = .68, DE = .15, p = .001$) y la condición 3 ($M = .90, DE = .13, p = .003$). Sin embargo, no se encontró una diferencia significativa en los puntajes de los índices de creencia e interacción.

Se calcularon a su vez los índices de sesgo a mi favor sobre la base de los computados por Čavojevová et al. (2018). El índice para los silogismos válidos se determinó restando la proporción de respuestas correctas en los silogismos pro-aborto de la proporción de respuestas correctas en los silogismos anti-aborto para cada participante. En el caso de los silogismos inválidos, el índice se calculó restando la proporción de respuestas correctas en los silogismos anti-aborto de la proporción de respuestas correctas en los silogismos pro-aborto. Para obtener un índice general del sesgo, se sumaron ambos índices. De este modo, cuanto más pronunciada fuera la actitud anti-aborto de los participantes, más alto debería ser el índice. Todos los índices mostraron medias negativas cercanas a cero, salvo en los índices inválidos de las condiciones 2 y 3, que fueron positivos, pero no significativamente diferentes de cero. Los índices con medias negativas significativamente mayores a cero fueron los índices global e inválido de la condición 1, y los índices válidos de las condiciones 2 y 3. Se realizó un ANOVA de medidas repetidas para comparar los índices entre las condiciones. Los índices globales no muestran diferencias significativas entre medias ($F_{(1, 69)} = 2.03, p = .150$). Los índices para los silogismos válidos tampoco difirieron significativamente ($F_{(1, 42)} = .47, p = .563$). En cambio, los índices para los silogismos inválidos mostraron diferencias significativas ($F_{(1, 44)} = 5.73, p = .013, \eta^2 = .149$). La condición 2 mostró el índice más alto ($M = .10, DE = .38$), diferenciándose de la condición 1 ($M = -.16, DE = .40, p = .002$). Es decir, la condición 2 mostró un mayor sesgo a mi favor al evaluar silogismos inválidos.

En cuanto a los puntajes globales de respuestas correctas, se llevó a cabo un ANOVA de medidas repetidas intra-sujeto para comparar las distintas condiciones. El análisis reveló un efecto principal ($F_{(1, 84)} = 8.53, p < .001, \eta^2 = .073$). Específicamente, la cantidad de respuestas correctas fue significativamente más baja en la condición 1 ($M = .67, DE = .47$), en comparación con la condición 2 ($M = .77, DE = .42, p = .001$) y la condición 3 ($M = .73, DE = .45, p = .03$), como puede observarse en la Figura 1. No se

Figura 1.

Media de aciertos totales por condición



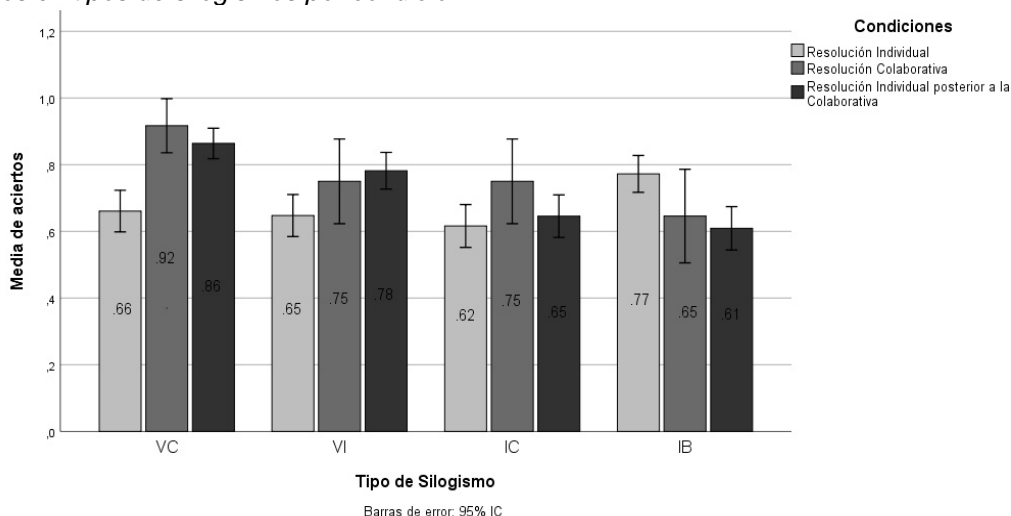
observaron diferencias significativas entre las últimas dos condiciones.

Además, se realizó otro ANOVA para evaluar los puntajes de corrección en cada tipo de silogismo. En particular, se respondieron correctamente más silogismos VC en la condición 2 ($F_{(1, 56)} = 29.6, p < .001, \eta^2 = .214, M = .92, DE = .28$), que en la condición 1 ($M = .66, DE = .47, p < .001$), y la condición 3 también se diferenció significativamente de la condición 1, mostrando más respuestas correctas ($M = .86, DE = .34, p < .001$). Además, se respondieron correctamente más silogismos VI en la condición 2 ($F_{(1, 81)} = 9.76, p < .001, \eta^2 = .082, M = .75, DE = .44$), que en la condición 1 ($M = .65, DE = .48, p = .004$), y la

condición 3 también tuvo más respuestas correctas que la condición 1 ($M = .78, DE = .41, p = .001$). En cuanto a los silogismos IC, hubo un mayor número de respuestas correctas en la condición 2 ($F_{(1, 84)} = 5.72, p = .005, \eta^2 = .050$), en comparación con las condiciones 1 ($M = .62, DE = .49, p = .016$) y 3 ($M = .65, DE = .48, p = .003$). Finalmente, en el caso de los silogismos II, hubo un mayor número de respuestas correctas en la condición 1 ($F_{(2, 68)} = 9.62, p < .001, \eta^2 = .081, M = .77, DE = .42$), que en la condición 2 ($M = .65, DE = .48, p = .001$) y en la 3 ($M = .61, DE = .49, p = .002$). Estos resultados se representan en la Figura 2.

Figura 2.

Media de aciertos en tipos de silogismos por condición



Nota. VC: válidos creíbles, VI: válidos increíbles, IC: inválidos creíbles, II: inválidos increíbles.

Los resultados respaldan la primera hipótesis planteada, según la cual el desempeño colaborativo en la resolución de tareas deductivas supera al desempeño individual. Los datos obtenidos muestran que los grupos colaborativos obtuvieron mejores resultados en términos de respuestas correctas en comparación con los participantes que realizaron la tarea de manera individual.

Durante la condición 2, resolución colaborativa, se conformaron 25 grupos distribuidos de la siguiente manera: cuatro grupos estuvieron compuestos por tres participantes, cinco grupos por cuatro participantes, y los restantes por cinco participantes. Ocho de los grupos tenían una composición ideológica mixta (32%) donde por lo menos un integrante se había posicionado contra el aborto, mientras que el resto (68%) tuvo composición homogénea pro-aborto.

Inicialmente, se investigaron las proporciones de respuestas correctas e incorrectas por grupo en cada silogismo, con el objetivo de determinar si el modelo de combinación social *la verdad gana* es aplicable al comportamiento grupal en la resolución de la tarea. Se tuvo en cuenta también el modelo de mayoría. Sin embargo, en el presente estudio, no se contempló el modelo de equiprobabilidad, ya que existen sólo dos opciones de respuesta posibles (válido o inválido),

lo que establece de por sí una probabilidad del 50% para cada opción.

Las respuestas posibles eran dos: acierto = A (respuesta correcta) o error = E (respuesta incorrecta). Respecto de los grupos de cinco integrantes, existen seis clasificaciones posibles: cinco aciertos y ningún error (5-0), cuatro aciertos y un error (4-1), tres aciertos y dos errores (3-2), dos aciertos y tres errores (2-3), un acierto y cuatro errores (1-4), ningún acierto y cinco errores (0-5). De manera similar, para los grupos compuestos por cuatro integrantes, las clasificaciones posibles son cinco: 4-0, 3-1, 2-2, 1-3 y 0-4; y para los grupos de tres integrantes son cuatro: 3-0, 2-1, 1-2 y 0-3. Cada grupo fue clasificado para cada uno de los ocho silogismos, obteniendo así un total de 200 respuestas grupales, de las cuales 153 fueron correctas (76.5%). En la [Tabla 1](#), se presentan las frecuencias de clasificación de grupo y de respuestas correctas e incorrectas.

De la totalidad de las respuestas correctas emitidas por los grupos, en 17 ocasiones (11%) un solo participante del grupo había respondido correctamente en la condición 1. De estas 17 respuestas, 7 (41%) resultaron ser correctas. Hubo 5 casos (3%) en donde el grupo respondió correctamente sin que ningún integrante haya respondido correctamente durante la fase de

Tabla 1.

Frecuencias de aciertos (A) y errores (E) de las distintas clasificaciones de grupos, para los ocho silogismos categorizados por validez y credibilidad

Tipo de grupo		Silogismo											
		VC			VI			IC			II		
		1	2	3	1	2	3	1	2	3	1	2	3
A	E	f	A	E	f	A	E	f	A	E	f	A	E
5	0	1	0	1	2	2	0	1	1	0	8	7	1
4	1	6	6	0	3	2	1	1	1	0	6	4	2
3	2	2	1	1	2	2	0	5	4	1	2	0	2
2	3	4	3	1	5	2	3	3	2	1	-	-	-
1	4	2	2	0	3	0	3	4	1	3	-	-	-
0	5	1	1	0	1	1	0	2	0	2	-	-	-
4	0	1	1	0	1	1	0	1	1	0	3	2	1
3	1	1	1	0	2	1	1	-	-	-	1	0	1
2	2	3	3	0	1	1	0	3	1	2	1	1	0
1	3	-	-	-	-	-	-	1	1	0	-	-	-
0	4	-	-	-	1	0	1	-	-	-	-	-	-
3	0	1	1	0	1	1	0	1	1	0	-	-	-
2	1	1	1	0	1	1	0	1	1	0	3	3	0
1	2	2	1	1	1	0	1	-	-	-	1	1	0
0	3	-	-	-	1	1	0	2	1	1	-	-	-

Continuación

		5			6			7			8		
5	0	5	5	0	5	5	0	3	3	0	4	3	1
4	1	4	4	0	6	6	0	9	7	2	4	2	2
3	2	5	5	0	2	2	0	3	2	1	4	2	2
2	3	1	1	0	2	1	1	1	1	0	3	0	3
1	4	-	-	-	1	1	0	-	-	-	1	0	1
0	5	1	1	0	-	-	-	-	-	-	-	-	-
4	0	1	1	0	2	2	0	2	2	0	2	2	0
3	1	3	3	0	3	3	0	2	2	0	2	1	1
2	2	1	1	0	-	-	-	1	1	0	-	-	-
1	3	-	-	-	-	-	-	-	-	-	1	0	1
0	4	-	-	-	-	-	-	-	-	-	-	-	-
3	0	2	2	0	2	2	0	2	2	0	2	2	0
2	1	2	2	0	2	1	1	2	2	0	2	2	0
1	2	-	-	-	-	-	-	-	-	-	-	-	-
0	3	-	-	-	-	-	-	-	-	-	-	-	-

Nota. VC: válidos creíbles, VI: válidos increíbles, IC: inválidos creíbles, II: inválidos increíbles, A: aciertos, E: errores.

resolución individual (condición 1), es decir, los casos de clasificación de grupo 0-5, 0-4 o 0-3. El modelo *la verdad respaldada gana* se cumple en un 20.91% de las respuestas grupales correctas, mientras que el modelo de mayoría se cumple en 48.37%.

La hipótesis 2, que sostiene que el modelo *la verdad gana* se aplica en la resolución grupal de la tarea deductiva, no se cumple en este estudio. Sin embargo, los resultados muestran que el modelo de mayoría desempeña un papel más significativo en la toma de decisiones grupal, representando un mayor porcentaje de respuestas correctas.

Discusión

Los resultados obtenidos son consistentes con la hipótesis 1, en tanto el desempeño colaborativo resulta superior al desempeño individual en la resolución de la tarea deductiva. En esta misma línea, a partir de los datos analizados, se observa que el índice de lógica es significativamente mayor en las condiciones 2 y 3, es decir, en la resolución colaborativa de la tarea y en la resolución individual que prosigue a la resolución colaborativa. De esta manera, se muestra que las respuestas grupales son más lógicas que las respuestas individuales y que las respuestas individuales se vuelven más lógicas luego de haber mediado una instancia de discusión grupal en la resolución de la tarea deductiva. Esto permite concluir que, en sintonía con estudios precedentes y con la hipótesis del origen dialógico

de la deducción, el desempeño deductivo humano es superior en entornos colaborativos.

Sin embargo, respecto de la hipótesis 1, es necesario plantear una excepción en relación con la resolución de los silogismos II. Para este tipo de silogismos, se ha observado que la resolución colaborativa no resulta superior a la resolución individual. Planteándose, de esta manera, una posible anomalía para las teorías que abogan a favor del origen social del razonamiento. Este aspecto del razonamiento deductivo, que involucra argumentos inválidos con conclusiones que no resultan creíbles para los razonadores, requeriría ser analizado con mayor precisión en investigaciones futuras. De todas maneras, es importante destacar que los estudios previos sobre razonamiento deductivo, en general, señalan que el razonamiento con argumentos inválidos supone mayores dificultades de resolución en comparación con los argumentos válidos (Čavojová et al., 2018; Evans et al., 1983).

En línea con esta observación, vale resaltar que el sesgo a mi favor se ve incrementado durante la evaluación colaborativa de silogismos inválidos. Así, puede verse que la actitud ideológica de las personas influyó especialmente en el razonamiento colectivo con silogismos inválidos. En este sentido, se podría hipotetizar que, para el caso de los argumentos generalmente considerados más difíciles de resolver, a saber, los argumentos inválidos, la instancia de razonamiento colectivo incrementó el sesgo a mi favor de los participantes.

Es importante señalar que la muestra de este estudio estuvo compuesta en su gran mayoría (91%) por participantes con una postura pro-aborto. Si se observa con mayor detalle la composición de los grupos, se evidencia que, en los grupos de composición mixta, es decir, grupos que contienen tanto participantes con posturas pro-aborto como anti-aborto, la resolución de la tarea arrojó un número significativamente menor de respuestas correctas en comparación con las respuestas dadas por los grupos con una composición ideológica homogénea. En el caso de los grupos mixtos, todos contaron con una mayoría de integrantes con posturas pro-aborto. Sería interesante replicar el estudio para analizar qué sucede con la resolución de esta tarea en grupos mixtos con mayoría de participantes con posturas anti-aborto, en grupos homogéneos con posiciones anti-aborto y en grupos con la misma cantidad de integrantes tanto con posturas pro como anti-aborto.

Respecto de los modelos de combinación social analizados, el modelo que mejor describe la toma de decisiones grupal en la resolución de la tarea deductiva es el modelo de mayoría. Estos resultados contradicen la hipótesis 2, así como también se diferencian de los datos aportados por estudios previos que muestran que el modelo que mejor describe la resolución de tareas intelectivas es el modelo *la verdad gana* (Laughlin, 2011). La idea subyacente al modelo *la verdad gana* sugiere que el miembro de un grupo que mejor comprende la lógica inherente a la tarea deductiva es capaz de convencer al resto de los miembros de su grupo de su punto de vista. Sin embargo, en nuestro estudio, no se observa este modelo de combinación social. Por el contrario, los resultados del presente estudio muestran que el modelo de mayoría es el que mejor describe la toma de decisiones colectiva. En general, el modelo de mayoría resulta el más adecuado para describir la resolución grupal de tareas de juicio, es decir, tareas evaluativas o juicios estéticos para los cuales no existe una única solución demostrablemente correcta (Laughlin, 2011). Podría especularse que, como la tarea deductiva empleada en este estudio involucra contenido ideológicamente controversial, como lo es la postura respecto del aborto, esto pudo haber influido en la manera de afrontar su resolución. Así, al tratarse de una tarea con contenido controvertido, la postura ideológica de la mayoría

de los integrantes del grupo podría haber jugado un rol central en la resolución de la tarea. Vale recordar que ninguno de los estudios antecedentes analizados incorpora en su discusión sobre el razonamiento colaborativo contenido ideológicamente controvertido. Esta posible explicación, sin embargo, es altamente especulativa y necesitaría ser respaldada por estudios a futuro diseñados específicamente para contrastar esta hipótesis.

Es importante señalar que se reportaron cinco instancias en las cuales ningún participante ofreció de manera individual la respuesta correcta; sin embargo, durante la fase de razonamiento colaborativo, los grupos integrados por esos participantes ofrecieron respuestas correctas. Este dato se alinea con las hipótesis que abogan a favor del origen dialógico del razonamiento humano, en tanto muestra que el intercambio argumentativo introduce mejoras en el desempeño deductivo humano. En sintonía con este hallazgo, en el estudio de Moshman y Geil (1998), también se reportaron casos en los cuales, si bien durante la fase de razonamiento individual ninguno de los miembros de un grupo resolvió la tarea de Wason correctamente, durante la fase de razonamiento colectivo el grupo llegó a consensuar la respuesta correcta.

El presente estudio no está exento de limitaciones. En primer lugar, el hecho de haber empleado un muestreo por conveniencia inevitablemente excluye a grandes segmentos de la población, restringiendo así la representatividad de la muestra. No obstante, al tratarse de un estudio de razonamiento deductivo, se hace bastante costoso en términos económicos acceder a muestras más representativas y brindarles las condiciones materiales adecuadas para la resolución de las tareas que luego se analizan. En segundo lugar, y también en relación con la muestra empleada, una consecuencia de este muestreo por conveniencia fue trabajar exclusivamente con población universitaria en su gran mayoría joven, femenina y con posturas pro-aborto. Un desafío pendiente para futuras investigaciones es conseguir mayor representatividad de género y de posturas ideológicas. En tercer lugar, otra limitación fue haber trabajado con grupos que no estuvieron conformados por el mismo número de participantes. Así, si bien la mayoría de los grupos estuvieron constituidos por cinco integrantes,

también hubo grupos de cuatro y de tres integrantes. Como el número de participantes puede tener efectos en la toma de decisiones colectivas, en futuros estudios sería preferible unificar la cantidad de miembros que integran cada grupo.

Respecto de futuras investigaciones, una línea promisoría consiste en analizar si la competencia intergrupala contribuye a mejorar el desempeño colaborativo, como se ha sugerido en un estudio reciente (Domberg et al., 2021). Profundizar en esta línea de investigación podría aportar información valiosa sobre los factores que inciden en el razonamiento colaborativo.

Disponibilidad de datos



Todo el conjunto de datos que apoya los resultados de este estudio fue publicado en OSF y puede ser accedido en <https://osf.io/jnmdf/>

Disponibilidad de métodos analíticos



Todo el conjunto de métodos analíticos que apoya los resultados de este estudio fue publicado en OSF y puede ser accedido en <https://osf.io/jnmdf/>

Disponibilidad de materiales



Todo el conjunto de materiales que apoya los resultados de este estudio fue publicado en OSF y se puede acceder en <https://osf.io/jnmdf/>

Conflictos de interés

Las autoras declaran no tener ningún conflicto de interés.

Referencias

American Psychological Association. (2020). *Manual de publicaciones de la American Psychological Association* (7ª ed.). Editorial El Manual Moderno.

Asociación Médica Mundial. (2024). *Declaración de Helsinki de la AMM – Principios éticos para las investigaciones médicas en seres humanos*. World Medical Association. <https://www.wma.net/es/policias-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos>

Béghin, G., & Markovits, H. (2024). The effects of strength and activation level of belief on belief-

biased reasoning. *Psychonomic Bulletin & Review*, 31(4), 1782–1788. <https://doi.org/10.3758/s13423-024-02461-z>

Boissin, E., Caparos, S., & De Neys, W. (2023). No easy fix for belief bias during syllogistic reasoning? *Journal of Cognitive Psychology*, 35(4), 401–421. <https://doi.org/10.1080/20445911.2023.2181734>

Boku, M., Yama, H., & Mercier, H. (2018). Robust improvements on reasoning performance following discussion in Japan. *Japanese Psychological Research*, 60(1), 47–53. <https://doi.org/10.1111/jpr.12180>

Castelain, T., Giroto, V., Jamet, F., & Mercier, H. (2016). Evidence for benefits of argumentation in a Mayan indigenous population. *Evolution and Human Behavior*, 37(5), 337–342. <https://doi.org/10.1016/j.evolhumbehav.2016.02.002>

Čavojová, V., Šrol, J., & Adamus, M. (2018). My point is valid, yours is not: myside bias in reasoning about abortion. *Journal of Cognitive Psychology*, 30(7), 656–669. <https://doi.org/10.1080/20445911.2018.1518961>

de Chantal, P., Newman, I., Thompson, V., & Markovits, H. (2020). Who resists belief-biased inferences? The role of individual differences in reasoning strategies, working memory, and attentional focus. *Memory & Cognition*, 48(4), 655–671. <https://doi.org/10.3758/s13421-019-00998-2>

Domberg, A., Tomasello, M., & Köymen, B. (2021). Collaborative reasoning in the context of group competition. *PLoS ONE*, 16(2), Artículo e0246589. <https://doi.org/10.1371/journal.pone.0246589>

Dutilh-Novaes, C. (2021). *The dialogical roots of deduction*. Cambridge University Press.

Evans, J. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, 128(6), 978–996. <https://doi.org/10.1037/0033-2909.128.6.978>

Evans, J. (2005). Deductive reasoning. En K. Holyoak & R. Morrison (Eds.) *The Cambridge Handbook of thinking and reasoning* (pp. 169–184). Cambridge University Press.

Evans, J. (2017). Belief bias in deductive reasoning. En R. F. Pohl (Ed.), *Cognitive illusions: Intriguing phenomena in thinking, judgment and memory* (2nd ed., pp. 165–181). Routledge/Taylor & Francis Group.

Evans, J., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, 11(4), 382–389. <https://doi.org/10.1080/13546780542000005>

Evans, J., Barston, J.L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, 11(3), 295–306. <https://doi.org/10.3758/BF03196976>

- Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.
- Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, 1(2), 134–148. [https://doi.org/10.1016/0010-0285\(70\)90009-5](https://doi.org/10.1016/0010-0285(70)90009-5)
- Köymen, B., O'Madagain, C., Domberg, A., & Tomasello, M. (2020). Young children's ability to produce valid and relevant counter-arguments. *Child Development*, 91(3), 685–693. <https://doi.org/10.1111/cdev.13338>
- Laughlin, P. R. (2011). *Group problem solving*. Princeton University Press
- Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellectual tasks. *Journal of Experimental Social Psychology*, 22(3), 177–189. [https://doi.org/10.1016/0022-1031\(86\)90022-3](https://doi.org/10.1016/0022-1031(86)90022-3)
- Levesque, H. J. (1986). Making believers out of computers. *Artificial Intelligence*, 30(1), 81–108. [https://doi.org/10.1016/0004-3702\(86\)90068-8](https://doi.org/10.1016/0004-3702(86)90068-8)
- Massolo, A., Traversi, M., & Alfonso, M. (2024). Belief bias in individual and collective reasoning. *Europe's Journal of Psychology*, 20(4), 317–328. <https://doi.org/10.5964/ejop.12041>
- Mercier, H., & Claidière, N. (2022). Does discussion make crowds any wiser? *Cognition*, 222, Artículo 104912. <https://doi.org/10.1016/j.cognition.2021.104912>
- Mercier, H., Deguchi, M., Van der Henst, J.-B., & Yama, H. (2016). The benefits of argumentation are cross-culturally robust: The case of Japan. *Thinking & Reasoning*, 22(1), 1–15. <https://doi.org/10.1080/13546783.2014.1002534>
- Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking & Reasoning*, 4(3), 231–248. <https://doi.org/10.1080/135467898394148>
- Over, D., & Evans, J. (2024). *Human reasoning*. Cambridge University Press.
- Stanovich, K. (2021). *The Bias that Divides Us. The Science and Politics of Myside Thinking*. MIT Press.
- Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, 94(4), 672–695. <https://doi.org/10.1037/0022-3514.94.4.672>
- Stanovich, K., & Toplak, M. (2023). Actively open-minded thinking and its measurement. *Journal of Intelligence*, 11(2), Artículo 27. <https://doi.org/10.3390/jintelligence11020027>
- Thompson, V. A., Striemer, C. L., Reikoff, R., Gunter, R. W., & Campbell, J. I. (2003). Syllogistic reasoning time: Disconfirmation disconfirmed. *Psychonomic Bulletin & Review*, 10(1), 184–189. <https://doi.org/10.3758/bf03196483>
- Tomasello, M. (2022). *The evolution of agency. Behavioral organization from lizards to humans*. MIT Press.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, 143(5), 1958–1971. <https://doi.org/10.1037/a0037099>
- Vedejová, D., & Čavojeová, V. (2021). Confirmation bias in information search, interpretation, and memory recall: evidence from reasoning about four controversial topics. *Thinking & Reasoning*, 28(1), 1–28. <https://doi.org/10.1080/13546783.2021.1891967>
- Wason, P. C. (1968). Reasoning about a Rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273–281. <https://doi.org/10.1080/14640746808400161>