

SESGOS DE RAZONAMIENTO, LENGUAJES FORMALES Y ENSEÑANZA DE LA LÓGICA

REASONING BIASES, FORMAL LANGUAGES AND TEACHING LOGIC

Alba Massolo*

Universidad Nacional de Córdoba,
Córdoba - Argentina.

Recibido mayo de 2019/Received May, 2019
Aceptado agosto de 2019/Accepted August, 2019

RESUMEN

Los estudios empíricos provenientes del ámbito de la psicología cognitiva han mostrado que el razonamiento humano no se guía por las prescripciones de la lógica. Por el contrario, estos estudios advierten que el razonamiento humano se apoya en la construcción de modelos mentales de baja demanda computacional y que existen ciertos patrones de procesamiento, denominados sesgos cognitivos, que están directamente vinculados a los típicos errores de razonamiento. El objetivo de este trabajo consiste en analizar de qué manera estas características del razonamiento humano están vinculadas con la aprehensión del concepto de consecuencia lógica. Para esto, realizamos un estudio exploratorio en un grupo de estudiantes universitarios de un curso introductorio de lógica de primer orden, a fin de observar si existe alguna relación entre el uso de lenguajes formales y la evaluación de argumentos. Encontramos que existe una correlación significativa entre el uso de lenguajes formales, por un lado, y de métodos sintácticos, por otro, y la capacidad para establecer la validez de un argumento. Estos resultados apoyarían la hipótesis de que los métodos y lenguajes formales constituyen una herramienta indispensable para la enseñanza de la lógica. Si bien las propuestas pedagógicas más cercanas al razonamiento intuitivo pueden ser útiles en una primera aproximación a conceptos centrales en lógica, nuestro estudio advierte algunas consecuencias negativas de estas propuestas didácticas que estarían vinculadas a la activación de algunos sesgos de razonamiento.

Palabras Clave: Sesgos cognitivos, consecuencia lógica, lenguajes formales, modelos mentales, enseñanza de la Lógica.

ABSTRACT

Empirical studies from the field of cognitive psychology have shown that human reasoning does not obey the prescriptions of logic. On the contrary, these studies showed that human reasoning relies on the construction of mental models and that there exist certain information processing patterns, called cognitive biases, which are linked to reasoning errors. The aim of this paper is to analyze if these human reasoning features have an impact on the apprehension of the concept of logical consequence. We conducted an exploratory study in a group of university students of an introductory course of Logic, in order to determine if there is any relation between formal languages and argument evaluation. As a result, our study showed that there exists a significant correlation between formal languages, on the one hand, and syntactic methods, on the other hand, and the ability for checking whether an argument is valid. These results would support the hypothesis that formal languages and formal methods are essential tools for teaching logic. Although pedagogical proposals that are close to naïve reasoning could be useful in the first approach to core concepts in logic, our study reveals some negative consequences related to these didactic proposals that are linked to reasoning biases activation.

Key Words: Cognitive biases, logical consequence, formal languages, mental models, teaching Logic.

* Autor correspondiente / Corresponding author: albamassolo@gmail.com

1. INTRODUCCIÓN

La relación entre la lógica y el razonamiento humano pasó de ser un fenómeno completamente incuestionable a ser un fenómeno ampliamente discutido. De hecho, a partir de la segunda mitad del pasado siglo se sostuvo, por un lado, que la lógica no tiene ninguna vinculación con el razonamiento humano (Harman, 2002) y, por otro lado, que el razonamiento humano no está determinado por las leyes de la lógica (Evans, 2005). En este contexto de discusión, diversas investigaciones han comenzado a analizar qué relación existe entre la lógica, entendida como disciplina, y el razonamiento humano (Woleński, 2016). Uno de los aspectos más debatidos ha sido la preponderancia de la lógica (o al menos de una teoría lógica en particular) como modelo normativo para el razonamiento humano (Stenning & van Lambalgen, 2008). A partir de algunas de las teorías desarrolladas en el marco de esta discusión, proponemos analizar de qué manera estas características del razonamiento humano influyen en el contexto de la enseñanza de la lógica. En particular, nuestro interés reside en indagar, a partir del análisis de un estudio exploratorio, cuál es la importancia del uso de lenguajes y métodos formales en la enseñanza de algunos conceptos centrales en lógica. Específicamente, centraremos nuestro análisis en los conceptos de consecuencia lógica y validez.

Para esto, en la sección II, caracterizamos las nociones de consecuencia lógica (Gómez-Torrente, 2009) y de sesgo cognitivo (Stanovich, West, & Toplak, 2016). A continuación, en la sección III, describimos uno de los núcleos teóricos centrales de la psicología del razonamiento: la teoría de los modelos mentales (Johnson-Laird, 2008) y desarrollamos una explicación sobre los sesgos cognitivos dada a partir de esta teoría. En la sección IV, presentamos una línea de investigación (Dutilh-Novaes, 2012) que argumenta a favor del uso de lenguajes y métodos formales en el contexto de la investigación científica. Posteriormente, en la sección V, exponemos nuestro estudio exploratorio y presentamos los principales resultados obtenidos. Finalmente, en la sección VI, analizamos los resultados de nuestro estudio a partir de la teoría de los modelos mentales, de la noción de sesgo de razonamiento y del uso de métodos y lenguajes formales. Esto nos permitirá elaborar algunas hipótesis explicativas que, por un lado, relacionan el uso de lenguajes interpretados

y de métodos semánticos con un alto porcentaje de error y con ciertas formas de razonamiento intuitivo vinculadas a los sesgos cognitivos. Y, por otro lado, relacionan el uso de lenguajes formales y métodos sintácticos con un mayor porcentaje de acierto y con formas de razonamiento normativamente correctas. De esta manera, nuestro estudio exploratorio nos permitirá hipotetizar a favor de ciertas ventajas pedagógicas tanto de los lenguajes formales como de los métodos sintácticos en el contexto de la enseñanza de la lógica.

2. LÓGICA, RAZONAMIENTO HUMANO Y SESGOS COGNITIVOS

Los trabajos sobre razonamiento humano realizados por Peter Wason a partir del año 1966 iniciaron una nueva tradición de investigación en el área de la psicología del razonamiento. En líneas generales, estos estudios empíricos consistieron en analizar de qué manera un grupo de personas resolvía alguna tarea de razonamiento. En estas tareas, los participantes debían elegir entre cuatro cartas cuál dar vuelta a fin de evaluar si una regla era verdadera o falsa (Wason, 1968) o, incluso, determinar si la conclusión de un argumento se seguía o no se seguía del conjunto de sus premisas (Evans, 2002). Los resultados mostraron que las respuestas dadas por los participantes diferían considerablemente de la respuesta esperada que se consideraba correcta. Así fue como se advirtió que existe una marcada divergencia entre el razonamiento humano y las leyes de la lógica (Perkins, 2002). Esta conclusión puso en jaque la, hasta entonces, indiscutida concepción de que el razonamiento humano se rige por las leyes de la lógica.

La consideración de una serie de estudios, tanto teóricos como empíricos, permite explicar, en gran medida, las razones de la divergencia entre la lógica, entendida como disciplina, y el razonamiento humano. Por un lado, la noción de consecuencia lógica definida para la lógica de primer orden estándar está caracterizada como una relación modal y formal (Gómez-Torrente, 2000). Esta definición, basada fundamentalmente en los trabajos de Tarski (1936), determina que la conclusión de un argumento es consecuencia lógica de sus premisas si y sólo si (i) toda interpretación que hace verdaderas a las premisas, también hace verdadera a la conclusión; es decir, que el argumento preserva necesariamente la verdad y (ii) la

condición (i) se cumple en virtud de la forma del argumento, es decir, que la preservación necesaria de la verdad responde exclusivamente a la forma lógica del argumento y no depende de ningún significado particular que puedan tener los conceptos incluidos en las premisas o en la conclusión. De esta manera, si un argumento es válido, todo argumento con su misma forma lógica también lo será. Los estudios experimentales comentados en el párrafo anterior consideraban que las tareas de razonamiento tenían una única respuesta correcta; esa respuesta era la indicada por la lógica de primer orden estándar¹.

Por otro lado, el análisis de los resultados empíricos comenzó a mostrar ciertos patrones en los errores cometidos en este tipo de tareas deductivas. Así, se advirtió que ciertos errores del razonamiento humano no son meramente errores fortuitos. Por el contrario, se trata de errores sistemáticos que se registran en un gran número de respuestas (Kahneman & Tversky, 1973). La sistematicidad de estos resultados llevó a denominar a estas respuestas alternativas *sesgos cognitivos*. Entre estos sesgos cognitivos, hay algunos que están exclusivamente vinculados al razonamiento y por eso se los denomina *sesgos de razonamiento* (Khemlani & Johnson-Laird, 2017). Uno de los sesgos de razonamiento más estudiados es el sesgo de la creencia (Evans, 2017; Klauer, Musch, & Naumer, 2000; Thompson & Evans, 2012), que señala la tendencia de los seres humanos a indicar que un argumento es válido sólo porque se cree que la conclusión de ese argumento es verdadera. Este sesgo muestra de qué manera el contexto y el contenido influyen en la evaluación de la calidad interna de un argumento. Como señalamos anteriormente, desde un punto de vista lógico, la validez de un argumento se determina exclusivamente en virtud de la forma del argumento y no depende del valor de verdad que de hecho pueda atribuirse a las premisas o a la conclusión. Sin embargo, el sesgo de la creencia advierte que la evaluación de un argumento puede verse afectada por las creencias previas de su evaluador y el contenido semántico del argumento. La investigación en torno a los sesgos cognitivos dio lugar a un debate entre quienes sostienen que los sesgos constituyen un problema que debería revertirse (Stanovich & West, 2000; Stanovich *et al.*, 2016) y quienes sostiene que los sesgos constituyen una respuesta óptima de adaptación al medio (Gigerenzer, 2008). Más

allá de este debate, que continúa en la actualidad, nuestro abordaje sobre los sesgos de razonamiento se enmarca en el ámbito de la enseñanza de la lógica y, por esta razón, los sesgos van a considerarse tendencias naturales de los estudiantes que entrañan un obstáculo en el aprendizaje de ciertos conceptos centrales de la lógica, como es el caso de la noción de consecuencia lógica. Esto no implica que los sesgos cognitivos no puedan considerarse ventajas adaptativas en otros contextos.

3. MODELOS MENTALES Y SESGOS DE RAZONAMIENTO

En el ámbito de la psicología del razonamiento, la teoría de los modelos mentales propuesta por Johnson-Laird (2001) ha permitido formular una explicación sobre el sesgo de la creencia ampliamente aceptada. Según esta teoría, los seres humanos razonan imaginando posibilidades, es decir, construyendo modelos mentales. Frente a un conjunto de premisas, los razonadores construyen una representación mental de la información dada. Luego, para establecer si una determinada conclusión se sigue o no se sigue de ese conjunto de premisas, imaginan situaciones posibles compatibles con esa representación mental. Así, si la conclusión es compatible con la situación representada, los razonadores consideran que el argumento es válido. En cambio, si la conclusión no es compatible con el escenario representado, concluyen que el argumento es inválido.

Los modelos mentales poseen diferencias significativas con respecto a los modelos formales de la semántica de primer orden. Johnson-Laird (2012) resume tres de los rasgos característicos de los modelos mentales, claramente ninguno de estos rasgos es compartido por la teoría de modelos de la semántica formal inspirada en los trabajos de Tarski. Estos principios son: el principio de iconicidad, el principio de posibilidad y el principio de verdad. En primer lugar, según el principio de iconicidad, la estructura de los modelos mentales es idéntica a la estructura que representa. Esto permitiría hacer un mapeo entre las partes y las relaciones del modelo mental y las partes y las relaciones de la escena descrita en determinado argumento. Una de las ventajas de la iconicidad es que posibilita arribar a conclusiones que no fueron explícitamente afirmadas en las premisas. Por ejemplo, en la descripción del plano de una ciudad podrían afirmarse las dos

siguientes premisas: “la iglesia está a la izquierda de la farmacia” y “la estación de la policía está a la derecha de la farmacia”. Así, se podría inferir que “la farmacia está entre la estación de policía y la iglesia” o que “la iglesia está a la derecha de la estación de policía”.

En segundo lugar, según el principio de posibilidad, cada modelo mental representa una sola posibilidad. Debido a las limitaciones de la memoria de trabajo, los razonadores pueden contemplar una única posibilidad a la vez. Esto produce dificultades al tratar de representar información que contiene disyunciones o condicionales, ya que cada oración compuesta por alguna de estas dos conectivas supone más de una alternativa posible. Sin embargo, contemplar todas estas posibilidades, como se hace en la semántica formal de asignaciones de verdad, implicaría una sobrecarga de la memoria de trabajo. Por ejemplo, para el caso de una disyunción inclusiva como “Lionel juega de media punta o de segundo delantero”, hay tres posibilidades que hacen verdadera esta afirmación: que Lionel juegue de media punta, que Lionel juegue de segundo delantero, o que Lionel juegue en ambas posiciones. Mantener estas tres posibilidades en la memoria de trabajo supone un gran esfuerzo para cualquier razonador. Sin embargo, para poder realizar una deducción correcta desde un punto de vista lógico deberían contemplarse las tres posibilidades a la vez. Esta segunda característica de los modelos mentales explica una gran parte de las diferencias entre el razonamiento humano y el razonamiento que aplica semánticas formales.

Por último, el principio de verdad postula que los modelos mentales representan sólo la información verdadera, ya sea que ésta se presente afirmada o negada; pero estos modelos nunca representan información falsa. Por ejemplo, frente a la oración “en la sala hay un sofá si y sólo si no hay una mesa”, las dos posibilidades que las personas generalmente representan al razonar son: una sala donde hay un sofá y una sala donde no hay una mesa. No obstante, las dos posibilidades que también se podrían representar a partir de esa oración, pero que son falsas, son que en la sala no hay un sofá y que en la sala hay una mesa. Esta estrategia cognitiva parece también direccionada a aliviar el funcionamiento de la memoria de trabajo, aunque claramente conlleva consecuencias indeseables que afectan la calidad del razonamiento.

Otro hallazgo que nos interesa resaltar de la teoría de los modelos mentales son los conceptos de modulación semántica y pragmática (Johnson-Laird & Byrne, 2002). El significado de las palabras que componen las oraciones de un argumento puede modular la interpretación de una conectiva. Por ejemplo, frente a las premisas “Lionel practica fútbol o Lionel practica algún deporte” y “Lionel no practica ningún deporte” el significado de la palabra “fútbol” bloqueará la construcción de un modelo mental que represente a Lionel practicando fútbol. No obstante, si razonamos siguiendo el silogismo disyuntivo, “Lionel practica fútbol” se sigue válidamente de las dos premisas anteriores. Similarmente, la cultura general o el conocimiento contextual pueden producir modulaciones pragmáticas. En este caso, frente a la premisa “si raspas un fósforo, entonces éste se enciende”, la conclusión “si raspas un fósforo que está completamente mojado, éste se enciende” sería bloqueada por casi cualquier razonador, ya que esta conclusión no sería representable como un modelo mental posible. Sin embargo, desde un punto de vista lógico, el argumento es válido.

La teoría de los modelos mentales sostiene que frente a una tarea de razonamiento que consiste en evaluar la validez de un argumento los razonadores seguirán los tres siguientes pasos: (1) construir una representación mental de la información dada en las premisas, (2) encontrar una conclusión que sea verdadera en esa representación y, finalmente, (3) tratar de encontrar una representación alternativa de las premisas y chequear que la conclusión también sea verdadera en esas representaciones. En caso de que esto no se dé, los razonadores procederán a chequear qué otras conclusiones podrían ser compatibles con esas representaciones alternativas. No obstante, si el razonador cree en la verdad de la conclusión, el tercer paso, es decir, la búsqueda de un contra-ejemplo, podría cancelarse. En cambio, si el razonador cree que la conclusión es falsa, es más probable que intente buscar una representación mental que haga verdaderas a las premisas y falsa a la conclusión (Garnharm & Oakhill, 2005). De esta manera, la teoría de los modelos mentales explica el sesgo de la creencia. Aunque ha sido mucho menos estudiado que el sesgo de la creencia, el sesgo focal (Toplak, West, & Stanovich, 2014) también constituye una marcada tendencia del razonamiento humano que conlleva perniciosas consecuencias en las tareas

de evaluación de argumentos. Este sesgo está vinculado con la avaricia cognitiva, es decir, la tendencia de los seres humanos a usar mecanismos de procesamiento de bajo costo computacional (Stanovich *et al.*, 2016). En el marco de la teoría de los procesos duales (Kahneman, 2011; Evans, 2012), se sostiene que los procesos cognitivos humanos se pueden diferenciar en dos tipos: procesos tipo 1, generalmente descriptos como automáticos, de bajos requerimiento computacional y que no hacen uso de la memoria de trabajo; y los procesos tipo 2 que son conscientes, tienen mayores requerimientos computacionales y necesitan de la memoria de trabajo. Si bien tradicionalmente los sesgos cognitivos fueron asociados con los procesos tipo 1, investigaciones más actuales han mostrado que los sesgos de razonamiento también pueden ser el resultado de procesos tipo 2 (Evans & Stanovich, 2013). De hecho, el sesgo focal está vinculado con los procesos cognitivos tipo 2. Cuando no es posible usar procesos tipo 1 y se requiere, en cambio, activar mecanismos cognitivos más complejos, el sesgo focal consiste en desplegar sólo algunas de las simulaciones cognitivas posibles o incluso en desplegar sólo una (Stanovich *et al.*, 2016). Así, por ejemplo, en relación con la teoría de los modelos mentales², el sesgo focal consistiría en la tendencia a representar sólo un modelo compatible con la información contenida en las premisas, o, de igual manera, en no representar todos los contra-modelos posibles para esa situación.

4. LENGUAJES FORMALES Y SESGOS DE RAZONAMIENTO

En el contexto del debate actual en torno a los sesgos cognitivos, Dutilh-Novaes (2012) argumentó a favor del uso de los lenguajes formales en el ámbito del descubrimiento científico y lógico-matemático. La autora estudió específicamente los efectos nocivos que el sesgo de la creencia podría tener en el descubrimiento científico. La influencia de las creencias previas puede resultar un obstáculo cuando se intenta obtener nueva información, cuestionar el conocimiento obtenido, construir una nueva explicación acerca de un fenómeno o incluso evaluar objetivamente la plausibilidad de ciertas conclusiones. Estas cuatro habilidades son fundamentales en el descubrimiento científico. A partir de estas consideraciones, Dutilh-Novaes sostiene que los lenguajes y métodos formales actúan

como mecanismos de-semantificadores capaces de evitar la activación semántica característica del sesgo de la creencia.

En líneas generales, un lenguaje formal podría caracterizarse como un lenguaje artificial compuesto por símbolos básicos y reglas estrictas de formación de fórmulas. Si bien hay diversas maneras de entender el adjetivo formal en el ámbito de la filosofía de la lógica (Cf. Dutilh-Novaes, 2011), sólo nos centraremos en describir dos sentidos de formal que permitirán caracterizar a los lenguajes formales como tecnologías cognitivas. Estos dos sentidos son lo formal entendido como de-semantificación y lo formal entendido como cálculo (Dutilh-Novaes, 2012). Por un lado, lo formal entendido como de-semantificación hace referencia a un lenguaje abstraído de cualquier tipo de contenido, es decir, un lenguaje sin significados. Esta caracterización alude a un lenguaje de símbolos en tanto esos símbolos constituyen marcas sin ningún tipo de contenido semántico. El concepto de proceso de de-semantificación fue acuñado por Krämer (2003) para indicar un proceso de abstracción con respecto a la materia y el contenido de un lenguaje. Por otro lado, lo formal entendido como cálculo hace referencia al procedimiento de seguir de manera mecánica las instrucciones o reglas de formación y transformación de un sistema de símbolos. Esta idea de formal como cálculo (o algoritmo) implica operar en el sistema sin emplear el ingenio, puesto que sólo se requiere seguir una serie de reglas formuladas explícitamente. El cálculo además es independiente de procesos aleatorios y se realiza en un plazo finito de tiempo. En consecuencia, las dos características centrales de los lenguajes formales son la abstracción total con respecto al contenido y el proseguimiento mecánico de una serie de reglas para operar con esos símbolos abstractos.

Así, la de-semantificación refiere a un modo de manipular un sistema de símbolos de manera tal que el significado sea parcial o completamente abstraído. Además, el modo mecánico de operar con un lenguaje de símbolos de-semantificado implica aplicar de manera estricta las reglas de transformación definidas para ese sistema. La tesis esgrimida por Dutilh-Novaes sostiene que estas dos características de los lenguajes formales, y los formalismos en general, permiten caracterizarlos como tecnologías cognitivas. Esto es, como dispositivos específicos que permiten mejorar e

incluso modificar los procesos cognitivos y de razonamiento de los seres humanos. Según esta autora, los lenguajes formales son tecnologías cognitivas que fueron desarrolladas para facilitar una serie de procesos cognitivos. La importancia de los lenguajes formales radica en que los mecanismos de activación semántica que se accionan de manera automática en determinados procesos de razonamiento pueden ser controlados por los mecanismos de-semantificadores característicos de este tipo de lenguajes. Es así como el uso de lenguajes formales, caracterizados como lenguajes de-semantificados, permite suprimir los patrones de activación semántica y las creencias previas que generalmente se activan durante los procesos de razonamiento.

De esta manera, trabajar con herramientas formales controlaría la interferencia de nuestras intuiciones y creencias previas en los procesos de razonamiento indispensables para construir ciencia. Por supuesto que este análisis no niega que los sesgos cognitivos puedan ser ventajosos en ciertos contextos cotidianos de interacción humana. No obstante, en el contexto del descubrimiento científico estos sesgos constituyen un importante obstáculo. Sobre la base de este interesante estudio, nuestro objetivo en las próximas dos secciones será mostrar que estas herramientas formales también tienen una importancia pedagógica fundamental en el ámbito de la enseñanza de la lógica.

Con respecto a la teoría de los modelos mentales, es menester señalar que esta tesis ofrece una explicación acerca del razonamiento cotidiano o intuitivo. En este sentido, la teoría de los modelos mentales es una teoría descriptiva sobre los procesos cognitivos involucrados en el razonamiento humano. Ahora bien, la teoría que se imparte en las clases de los cursos introductorios a la lógica, generalmente la lógica clásica de primer orden, es una teoría normativa, esto es, una teoría que prescribe cómo se debe operar frente a un argumento. Advirtiendo esta diferencia, el objetivo de nuestro estudio exploratorio es doble: por un lado, analizar si ciertos errores típicos en la resolución de tareas vinculadas a las nociones de consecuencia lógica y validez podrían explicarse a partir de algunas de las características del razonamiento humano descritas por la teoría de los modelos mentales. Y, por otro lado, analizar si el uso de lenguajes formales podría correlacionarse con una mejora en el desempeño para resolver este tipo de tareas.

A partir de estos dos objetivos, se podrían plantear algunas hipótesis acerca de la conexión entre ambos desarrollos teóricos. En particular, analizar si el uso de lenguajes y métodos formales en el contexto de la enseñanza de la lógica podría controlar la activación de ciertos mecanismos vinculados a los sesgos de razonamiento descritos por la teoría de los modelos mentales. En este sentido, la tesis sobre el rol central de los lenguajes formales en el ámbito del descubrimiento científico podría extenderse al ámbito de la enseñanza de la lógica. De este modo, frente a ciertas predicciones típicas de la teoría de los modelos mentales, como el sesgo de la creencia y el sesgo focal, los lenguajes formales, en tanto tecnologías cognitivas, podrían contribuir a disminuir las diferencias existentes entre las prescripciones de la lógica clásica de primer orden, por un lado, y la descripción sobre el razonamiento humano ofrecida por la teoría de los modelos mentales, por otro.

5. ESTUDIO EXPLORATORIO

Realizamos un estudio exploratorio en un grupo de estudiantes universitarios de un curso introductorio de lógica de primer orden a fin de analizar la relación entre los lenguajes formales, los métodos empleados y la capacidad para determinar la validez de un argumento. Los participantes asistían a un curso que usó como bibliografía una versión en español del libro de Barwise y Etchemendy (2002) que contiene un software, *el mundo de Tarski*. En este libro, se introduce un lenguaje de primer orden interpretado y se muestra cómo construir modelos (los mundos del *mundo de Tarski*) para chequear la validez de los argumentos. Asimismo, el libro presenta el método de deducción natural estilo Fitch para las pruebas formales y también muestra cómo construir demostraciones informales, es decir, pruebas que usan las reglas deductivas del sistema pero se escriben en lenguaje coloquial. Para el diseño de la tarea, se tuvo en cuenta la familiaridad de los participantes con el lenguaje y los métodos del *mundo de Tarski*.

Sobre este punto, es necesario señalar que el enfoque del *mundo de Tarski* recoge algunas de las críticas de Etchemendy (1990) al concepto tarskiano de consecuencia lógica descrito en la sección II. De acuerdo con Etchemendy, en la teoría formal de modelos existe una diferencia entre una semántica representativa, que representa

posibilidades, y una semántica interpretativa, que interpreta los ítems lingüísticos de las oraciones. Según Etchemendy, para dar cuenta adecuadamente del concepto de consecuencia lógica es necesario apelar a una semántica representativa. El trabajo de Tarski, sin embargo, estaría basado en una semántica interpretativa. Atendiendo a esta crítica, Barwise y Etchemendy (2002) se inclinaron por una semántica representativa para construir *el mundo de Tarski*.

5.1. Método

5.1.2. Participantes

Participaron de este estudio exploratorio de manera voluntaria 45 estudiantes (15 mujeres, 29 varones y 1 participante que indicó pertenecer a otro género) que asistían a un curso introductorio de lógica de primer orden de la carrera de filosofía de una universidad nacional de Argentina. La edad media de la muestra es de 22,45 años (D. E.=7,06), mientras que su valor modal es de 18 años.

5.1.3. Tarea

La tarea consistió en indicar si un argumento compuesto por dos premisas y una conclusión era válido o inválido, es decir, si la conclusión era consecuencia lógica de las premisas. Se pidió además a los participantes que indicaran por escrito de qué manera habían determinado si la conclu-

establecer si un argumento es válido o inválido. También puede chequearse validez mediante el sistema de Fitch. Sin embargo, este método, que no es mecánico, sólo permite probar la validez de un argumento, pero no permite mostrar invalidez. Por esta razón, sólo se usaron argumentos válidos en nuestro estudio exploratorio.

Sobre la base de esta tarea, se elaboraron cuatro test diferentes: el Test 1 contenía un argumento válido expresado en un lenguaje de primer orden interpretado como el usado en el software *el mundo de Tarski*. El argumento de este primer test tenía premisas inconsistentes. El Test 2 contenía el mismo argumento del Test 1 pero expresado en un lenguaje de primer orden sin interpretación, es decir, en un lenguaje formal de primer orden. El Test 3 contenía un argumento válido diferente expresado en un lenguaje interpretado como en el usado en *el mundo de Tarski*. A diferencia del test 1, las premisas del argumento de este tercer eran consistentes. Finalmente, el Test 4 contenía el mismo argumento del Test 3 pero expresado en un lenguaje de primer orden formal (no interpretado) (ver tabla 1).

Tabla 1.

Los cuatro Test usados en el estudio exploratorio

Test	Argumento
Test 1	Premisas: $\text{Dodec}(c) \rightarrow \text{MayorQ}(a,b)$; $\text{Dodec}(c) \rightarrow \neg \text{MayorQ}(a,b)$ Conclusión: $\text{Dodec}(c) \rightarrow \text{Chico}(b)$
Test 2	Premisas: $A \rightarrow B$; $A \rightarrow \neg B$ Conclusión: $A \rightarrow C$
Test 3	Premisas: $(\text{Dodec}(c) \vee \text{Chico}(d)) \rightarrow \text{MayorQ}(a,b)$; $\text{Dodec}(c) \wedge \text{Chico}(d)$ Conclusión: $\text{MayorQ}(a,b)$
Test 4	Premisas: $(A \vee B) \rightarrow C$, $A \wedge B$ Conclusión: C

sión del argumento era (o no era) consecuencia lógica de las premisas, es decir, debían explicar qué método habían empleado en la resolución de la tarea. Para el fragmento de la lógica de primer orden que utilizamos en el estudio, una manera de determinar la validez de un argumento es a partir de tablas de verdad. Este método tiene la ventaja de permitir determinar de manera mecánica si un argumento es válido o inválido. Además, el libro de Barwise y Etchemendy muestra cómo chequear validez a partir de la construcción de mundos. Este segundo método no es mecánico pero permite

5.1.4. Procedimiento

A cada participante se le entregó una hoja que contenía uno de los cuatro test junto con las consignas de resolución de la tarea, de esta manera, cada participante resolvió, usando papel y lápiz, un único test. La distribución de cada test fue sistemática, intentando, de esta manera, que cada test sea resuelto por la misma cantidad de participantes. Cada participante resolvió el test de manera individual y no se permitieron realizar consultas o preguntas aclaratorias durante su resolución.

5.1.5. Análisis y resultados

Antes de comenzar con el análisis de los resultados, vamos a hacer algunas aclaraciones en relación con la categorización elegida para las variables. En primer lugar, con respecto al test resuelto, vamos a considerar, por un lado, el lenguaje del argumento. De esta manera, consideraremos de manera conjunta las respuestas dadas al Test 2 y al Test 4 y denominaremos a esta variable Test Lenguaje Formal (TLF); similarmente, las respuestas dadas al Test 1 y al Test 3 se unificarán bajo la variable Test Lenguaje Interpretado (TLI). Por otro lado, y a fin de analizar si existen diferencias con respecto al argumento con premisas inconsistentes y al argumento con premisas consistentes, también analizaremos los resultados manteniendo cada uno de los test como cuatro variables diferentes.

En segundo lugar, con respecto al método elegido por los participantes para determinar si el argumento era (o no era) válido, a partir de las respuestas obtenidas, hemos dividido la variable método en cuatro categorías: (1) método sintáctico, indica la construcción de una demostración formal en el sistema de deducción natural estilo Fitch o una demostración informal usando las reglas deductivas del sistema (2) método semántico, indica la construcción de mundos o el uso de tablas de verdad (únicamente 2 participantes mencionaron el uso de tablas de verdad); (3) método híbrido, para el caso de las respuestas que combinan tanto recursos semánticos como sintácticos; y por último, (4) método indeterminado, cuando no fue posible determinar claramente qué estrategia de justificación siguió el participante (Por ejemplo: casos que sólo se limitaron a responder “porque la conclusión es consecuencia lógica de las premisas”). Finalmente, queremos señalar que todo el análisis estadístico fue realizado usando el software *InfoStat*.

La tabla 2 muestra la relación entre el uso de lenguajes formales y el acierto al determinar la validez del argumento. Existe una correlación moderada ($\phi=0.33$) entre ambas variables. Del total de participantes que resolvió el Test Lenguaje Interpretado ($n=21$), sólo un 38,09% respondió correctamente que el argumento era válido, mientras que entre el total de participantes que resolvió el Test Lenguaje Formal ($n=24$) el porcentaje de respuestas correctas asciende al 70.83%. La tabla de frecuencias relativas muestra que el 17,78% del total de los participantes ($n=45$) resolvió correctamente el Test Lenguaje Interpretado; mientras que el Test Lenguaje Formal fue resuelto correctamente por un 37.78% del total de los participantes. Ahora bien, si analizamos la variable Test distinguiendo las cuatro categorías que corresponden a cada uno de los test resueltos por los participantes (ver tabla 3) observamos que esta correlación entre lenguajes formales y respuestas correctas se acentúa ligeramente ($\phi=0.39$) para el caso del argumento con premisas inconsistentes (Test 1 y Test 2). Así, del total de participantes que resolvieron estos dos test ($n=25$), entre quienes resolvieron el test expresado en lenguaje formal ($n=14$), un 57,14% respondió correctamente que el argumento era válido, mientras que sólo un 18% del total de participantes que resolvió el test con premisas inconsistentes en lenguaje interpretado ($n=11$) respondió correctamente. Además, con respecto a los métodos empleados para la resolución de ambos test, pudo observarse una correlación significativa (Coef. Conting. Cramer = 0.54) entre el uso de métodos sintácticos y el acierto en la respuesta acerca de la validez del argumento. En esta relación de variables, del total de participantes que empleó métodos sintácticos ($n=9$), el 88.88% indicó correctamente que el argumento era válido. Del total de participantes que empleó métodos semánticos ($n=11$), sólo el 18.18% indicó correctamente que el argumento era válido.

Tabla 2.

Relación Lenguaje del Test y respuesta correcta

Respuesta	Test Lenguaje Formal	Test Lenguaje Interpretado	Total
Válido	17	8	25
	37,78%	17,78%	55,56%
Inválido	7	13	20
	15,56%	28,89%	44,44%
Total	24	21	45
	53,33%	46,67%	100,00%

Tabla 3.
Relación Test y respuesta correcta

Respuesta	Test 1	Test 2	Test 3	Test 4	Total
Válido	2	8	6	9	25
	4,44%	17,78%	13,33%	20,00%	55,56%
Inválido	9	6	4	1	20
	20,00%	13,33%	8,89%	2,22%	44,44%
Total	11	14	10	10	45
	24,44%	31,11%	22,22%	22,22%	100,00%

Tabla 4.
Relación Lenguaje del Test y método de resolución

Test	Método 1	Método 2	Método 3	Método 4	Total
Lenguaje Formal	12	6	2	4	24
	26,67%	13,33%	4,44%	8,89%	53,33%
Lenguaje Interpretado	2	16	1	2	21
	4,44%	35,56%	2,22%	4,44%	46,67%
Total	14	22	3	6	45
	31,11%	48,89%	6,67%	13,33%	100,00%

La tabla 4 muestra la relación entre el lenguaje usado en el test y el método que utilizaron los participantes para determinar la validez del argumento. Existe una correlación moderada (Coef. Conting. Cramer =0.37) entre ambas variables. Esto indica que existe cierta preferencia para determinar la validez de un argumento a partir de métodos sintácticos en el caso de argumentos expresados en un lenguaje formal de primer orden. Entre quienes resolvieron el Test Lenguaje Formal (n=24), un 50% empleó métodos sintácticos y un 25% métodos semánticos; mientras que entre quienes resolvieron el Test Lenguaje Interpretado (n=21), un 76,19% empleó métodos semánticos y sólo un 9,52% eligió resolver el ejercicio usando métodos sintácticos.

Con respecto a la relación entre las variables método y respuesta correcta (ver tabla 5), encontramos que existe una correlación significativa entre estas dos variables (Coef. Conting. Cramer

=0.43). El método preferido por la mayoría de los participantes para determinar validez fue el método semántico (48,89%), seguido por el método sintáctico (31,11%). Sin embargo, mientras que el 92,85% del total de participantes que empleó métodos sintácticos en la resolución de la tarea (n=14) acertó; el porcentaje del total de participantes que respondió correctamente usando métodos semánticos (n=22) desciende al 45,45%. Por otra parte, la cantidad de aciertos de los casos donde no se pudo determinar qué método había sido empleado (n= 6) fue nula. Esto muestra una correlación significativa entre el empleo de métodos sintácticos y la capacidad para determinar correctamente la validez de un argumento.

Tabla 5.
Relación Método de resolución y respuesta correcta

Respuesta	Método 1	Método 2	Método 3	Método 4	Total
Válido	13	10	2	0	25
	28,89%	22,22%	4,44%	0,00%	55,56%
Inválido	1	12	1	6	20
	2,22%	26,67%	2,22%	13,33%	44,44%
Total	14	22	3	6	45
	31,11%	48,89%	6,67%	13,33%	100,00%

Tabla 6.
Relación Lenguaje del Test-Respuesta
correcta y método de resolución

Método	TLF – Válido	TLF- Inválido	TLI- Válido	TLI- Inválido	Total
Método 1	11 24,44%	1 2,22%	2 4,44%	0 0,00%	14 31,11%
Método 2	4 8,89%	2 4,44%	6 13,33%	10 22,22%	22 48,89%
Método 3	2 4,44%	0 0,00%	0 0,00%	1 2,22%	3 6,67%
Método 4	0 0,00%	4 8,89%	0 0,00%	2 4,44%	6 13,33%
Total	17 37,78%	7 15,56%	8 17,78%	13 28,89%	45 100,00%

En la tabla 6, mostramos la relación entre las variables test y respuesta correcta, por un lado, y la variable método, por otro. Los resultados indican que se trata de una correlación significativa (Coef. Conting. Cramer = 0.42). Asimismo, esta correlación aumenta levemente (Coef. Conting. Cramer = 0.50) con respecto al argumento con premisas inconsistentes. Finalmente, queremos señalar que el 55.56% del total de la muestra resolvió correctamente la tarea.

5. DISCUSIÓN GENERAL

En primer lugar, señalaremos algunos resultados del estudio exploratorio vinculados directamente con la teoría de los modelos mentales. Para esto, volvemos a enfatizar que los métodos semánticos, particularmente la construcción de modelos, fue la metodología más elegida (48.89%) por los participantes para resolver la tarea. La marcada preferencia por el empleo de métodos semánticos, basados en la construcción de mundos compatibles con la información obtenida a partir de las premisas, podría reflejar la familiaridad que encuentran los participantes entre esta forma de determinar la validez de un argumento y las formas de razonamiento más intuitivas. Estas serían las formas de razonar típicamente empleadas por poblaciones sin entrenamiento en lógica, basadas en la construcción de modelos mentales. Esto muestra cierta cercanía entre el enfoque representacional propio de la semántica del *mundo de Tarski* y los modelos mentales propuestos por Johnson-Laird.

No obstante, los principios de iconicidad, posibilidad y verdad descriptos por Johnson-Laird para diferenciar los modelos mentales de los modelos de la semántica formal podrían servir para explicar el elevado porcentaje de error entre los participantes que optaron por determinar la validez del argumento a partir del empleo de métodos semánticos. En la sección anterior señalábamos que el 45.45% de los participantes que usaron métodos semánticos resolvió correctamente la tarea. Esto muestra un pronunciado porcentaje de error (54.55%) en el razonamiento a partir de modelos formales. Una posible hipótesis explicativa para este fenómeno podría darse a partir de la teoría de los modelos mentales y el sesgo focal. En primer lugar, la consigna, junto con el contexto de resolución de la tarea, daría la pista de que se trata de una tarea cuya resolución requiere la implementación de procesos tipo 2. Esto descartaría casi de plano la activación de procesos cognitivos tipo 1. Sin embargo, como fue señalado en la sección III, el procesamiento de la información a partir de mecanismos tipo 2 no garantiza que se obtenga la respuesta correcta. En nuestro estudio, para resolver correctamente la tarea, se debían considerar todos los modelos posibles donde las premisas son verdaderas y chequear que en cada uno de esos modelos la conclusión también fuera verdadera. Al menos esto es lo que indica la teoría de modelos para la lógica de primer orden estándar. Sin embargo, se puede advertir que este proceso cognitivo supone una importante demanda para la memoria de trabajo. Por esto, la avaricia cognitiva podría desempeñar

un rol importante en la resolución de esta tarea. Así, en vez de desplegar todos los modelos posibles, se trabaja con la representación (ya sea mental o en lápiz y papel) de sólo algunos de esos modelos posibles, o incluso de un solo modelo. De esta manera, el sesgo focal podría explicar la tendencia a considerar menos representaciones de las que se debería considerar. Aquí residiría el error que llevaría a resolver de manera incorrecta la tarea. Claramente, esta hipótesis debería testearse con tareas que contengan argumentos inválidos, ya que típicamente el sesgo focal llevaría a considerar erróneamente que un argumento es válido porque no se construyó la representación mental del contra-modelo que lo refuta.

Una consideración especial merece el análisis del razonamiento a partir de premisas inconsistentes (Test 1 y Test 2). Recordemos que sólo un 18% del total de participantes que resolvió el test con premisas inconsistentes en lenguaje interpretado ($n=11$) respondió correctamente sobre la validez del argumento. En la lógica de primer orden estándar, el *ex falso quodlibet* indica que a partir de una contradicción se sigue cualquier conclusión. Si bien esta regla ha despertado una fuerte controversia en filosofía de la lógica (Cf. Carnielli & Coniglio, 2016), se trata de una regla válida en el sistema lógico que estamos considerando. Según la explicación del razonamiento humano dada por la teoría de los modelos mentales, la conjunción de dos modelos que contienen una proposición y la negación de esa proposición lleva a un modelo vacío, es decir, a una contradicción (Johnson-Laird, 2008). Así, como los razonadores buscan que la conclusión sea compatible con las representaciones de las premisas, al no poder construir modelos mentales compatibles con la representación de las premisas, tenderían a concluir que nada puede seguirse de tales modelos, puesto que, ninguna representación que pueda construirse es compatible con esa conclusión. Y de esta manera, justificarían erróneamente la invalidez del argumento. Esto podría explicar el alto porcentaje de error entre quienes optaron por el empleo de métodos semánticos para responder sobre la validez del argumento con premisas inconsistentes

En segundo lugar, señalaremos algunos resultados del estudio exploratorio vinculados con la tesis propuesta por Dutilh-Novaes sobre la importancia de los lenguajes formales. Indicamos, además, algunas vinculaciones que podrían

establecerse entre esta tesis y la teoría de los modelos mentales. En este sentido, con respecto a la relación entre el lenguaje en el que está expresado el argumento y la respuesta correcta acerca de la validez del mismo, vimos en la sección anterior que existe una moderada correlación ($\phi = 0.33$) entre ambas variables. Asimismo, también existe una correlación moderada entre el lenguaje del argumento y el método (Coef. Conting. Cramer = 0.37) y entre el método elegido y los aciertos al determinar la validez del argumento (Coef. Conting. Cramer = 0.43). A partir de estos datos, podríamos hipotetizar que la relación entre el lenguaje en el que está expresado el argumento y el porcentaje de aciertos está mediada por el uso de los métodos para chequear validez. Así, el lenguaje tendría cierta influencia sobre la elección del método: si un argumento está expresado en un lenguaje formal, habría una preferencia por usar métodos sintácticos para determinar su validez; mientras que, si el argumento está expresado en un lenguaje interpretado, la preferencia se inclinaría por usar métodos semánticos. A su vez, dada la correlación significativa entre métodos sintácticos y aciertos, por un lado, y métodos semánticos y respuestas erróneas, por otro, esto podría explicar el mejor desempeño encontrado entre los participantes que resolvieron el Test Lenguaje Formal frente a los participantes que resolvieron el Test Lenguaje Interpretado.

De esta manera, los lenguajes interpretados llevarían a preferir el uso de métodos semánticos. No obstante, como señalamos anteriormente, la avaricia cognitiva, reflejada principalmente en el sesgo focal, y el contenido semántico de las expresiones del argumento, llevarían a los participantes a razonar a partir de la construcción de modelos mentales. Si bien estos modelos mentales poseen ciertas semejanzas con los modelos formales de la semántica de primer orden, sus diferencias son esenciales. Esto explicaría el pronunciado porcentaje de error al resolver la tarea usando métodos semánticos. En cambio, los lenguajes formales llevarían a preferir el uso de métodos sintácticos. En este caso, al no haber contenido semántico ni pragmático involucrado, la apelación a formas de razonamiento basadas en la construcción de modelos se hace más improbable porque no hay contenido explícito para representar. Esto explicaría la preferencia por usar reglas deductivas, ya sea en una demostración formal o en una demostración

informal. Por lo tanto, el uso de lenguajes formales estimularía el uso de métodos sintácticos y de esta manera se evitarían las formas más cotidianas e intuitivas de razonar, directamente vinculadas con los sesgos de razonamiento expuestos en la sección II.

Con respecto al análisis del razonamiento a partir de premisas inconsistentes (Test 1 y Test 2), según señalamos en la sección V, el coeficiente de correlación más significativo de nuestro estudio (Coef. Conting. Cramer = 0.54) se da entre la metodología empleada para determinar la validez del argumento con premisas inconsistentes y el acierto al responder que se trata de un argumento válido. La hipótesis que construimos a partir del análisis de los datos indica que la relación entre el lenguaje del argumento y el acierto al determinar la validez del mismo está mediada por el método elegido. Así, por un lado, en el caso de los métodos semánticos, el contenido semántico de los lenguajes interpretados podría obstaculizar la resolución de la tarea. En este sentido, entraría en juego el concepto de modulación semántica. Así, en el caso del argumento del test 1, podríamos afirmar que la presencia de premisas inconsistentes bloquea la construcción de los modelos que confirmarían o refutarían la validez del argumento. En este sentido, se reflejarían claramente los efectos nocivos del lenguaje interpretado y de los métodos semánticos de construcción de mundos en la evaluación de un argumento. Por otro lado, en el caso de los métodos sintácticos, los lenguajes formales podrían bloquear la activación de maneras más intuitivas de razonamiento. El test 2, expresado en un lenguaje formal, fue resuelto de manera correcta por un 57,14% del total de los participantes que resolvió este test (n=14). A este respecto, se podría aducir, a partir de nuestra hipótesis, que el uso de lenguajes formales y de métodos sintácticos evitarían en gran medida los efectos nocivos del contenido semántico sobre los procesos de razonamiento. Por un lado, cuando un argumento se expresa en un lenguaje formal, su contenido semántico está reducido a un mínimo. Por otro lado, el razonamiento a partir de reglas posibilitaría realizar inferencias a partir de un lenguaje formal, de-semantificado. De esta manera, el uso de lenguajes formales y métodos sintácticos bloquearía la activación de sesgos de razonamiento vinculados al contenido semántico y de esta manera se reduciría considerablemente el porcentaje de error en el razonamiento.

De este modo, y en consonancia con la línea de investigación propuesta por Dutilh-Novaes, nuestro estudio exploratorio permitiría elaborar hipótesis sobre algunas ventajas pedagógicas tanto de los lenguajes formales como de los métodos sintácticos en el contexto de la enseñanza de la lógica. En este sentido, estos resultados alentarían las didácticas que enfatizan el trabajo con lenguajes formales y métodos sintácticos en lógica. Con respecto a la semántica, si bien en *el mundo de Tarski* se propone, a partir de la noción de mundo, una manera ingeniosa e intuitiva para explicar el concepto de modelo formal, nuestro estudio reveló algunos problemas derivados de esta aproximación a la semántica formal. En este sentido, nuestros resultados indicarían que el contenido de la semántica formal no debería limitarse al trabajo con oraciones y mundos construidos en *el mundo de Tarski*. Por el contrario, resulta necesario incluir también referencias a lenguajes y modelos formales de primer orden no interpretados. Vale señalar que nuestro estudio exploratorio se limitó a analizar el desempeño de un grupo de estudiantes a la hora de determinar la validez de un argumento usando lápiz y papel. Sería interesante analizar, para el caso de los test que emplean lenguajes interpretados, si la resolución de esta tarea usando el software *el mundo de Tarski* lleva a resultados similares o si, por el contrario, existen diferencias con respecto al desempeño usando lápiz y papel.

Finalmente, queremos volver a subrayar que se trata de un estudio exploratorio que nos permitió formular algunas hipótesis sobre la relación entre el razonamiento humano y la enseñanza de la lógica; por esto, resultaría necesario repetir este estudio en muestras más grandes para avanzar en el testeo de las hipótesis aquí sugeridas. El carácter exploratorio del estudio en conjunción con los coeficientes de correlación empleados, no permiten establecer relaciones de asociación causal entre las variables analizadas. De hecho, nuestro objetivo se limitó a analizar si existen relaciones significativas entre las variables elegidas. Asimismo, como el lector habrá podido advertir, las correlaciones que manejamos, si bien son significativas, en términos estadísticos se trata de correlaciones moderadas (sólo uno de los indicadores de correlación obtenidos supera 0.50). Este hecho nos obliga a enfatizar la necesidad de seguir refinando las hipótesis planteadas.

Referencias

- Barwise, J. & Etchemendy, J. (2002). *Language, Proof and Logic*. Stanford, U.S.A.: CSLI Publications.
- Carnielli, W. & Coniglio, M. (2016). *Paraconsistent Logic: Consistency, Contradiction and Negation*. Basel, Switzerland: Springer.
- Dutilh-Novaes, C. (2011). The Different Ways in which Logic is (Said to Be) Formal. *History and Philosophy of Logic*, 32(4), 303–332.
- Dutilh-Novaes, C. (2012). *Formal Languages in Logic*. Cambridge, U.K.: Cambridge University Press.
- Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge (MA), U.S.A.: Harvard University Press.
- Evans, J. St. B. T. (2002). Logic and Human Reasoning: an Assessment of the Deduction Paradigm. *Psychological Bulletin*, 128(6), 978–996.
- Evans, J. St. B. T. (2005). Deductive Reasoning. En K. Holyoak & R. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 169–184). Cambridge, U.K.: Cambridge University Press.
- Evans, J. St. B. T. (2012). Dual-Process Theories of Deductive Reasoning: Facts and Fallacies. En K. Holyoak & R. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 115–133). New York, U.S.A.: Oxford University Press.
- Evans, J. St. B. T. (2017). Belief Bias in Deductive Reasoning. En R. F. Polh (Ed.), *Cognitive Illusions: Intriguing Phenomena in Thinking, Judgment and Memory* (pp. 165–181). New York, U. S. A.: Routledge.
- Evans, J. St. B. T. & Stanovich, K. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241.
- Garnharm, A. & Oakhill, J. (2005). Accounting of Belief Bias in Mental Model Framework: Comment on Klauer, Musch, and Naumer (2000). *Psychological Review*, 112(2), 509–518.
- Gigerenzer, G. (2008). Why Heuristics Work. *Perspectives on Psychological Science*, 3(1), 20–29.
- Gómez Torrente, M. (2000). *Forma y modalidad. Una introducción al concepto de consecuencia lógica*. Buenos Aires, Argentina: Editorial Eudeba.
- Gómez Torrente, M. (2009). Rereading Tarski on Logical Consequence. *Review of Symbolic Logic*, 2(2), 249–297.
- Harman, G. (2002). Internal Critique: a Logic is not a Theory of Reasoning and a Theory of Reasoning is not a Logic. En D. Gabbay, H. J. Ohlbach, R. Johnson, & J. Woods (Eds.), *Handbook of the Logic of Argument and Inference* (pp. 171–186). North Holland, Netherlands: Elsevier.
- Johnson-Laird, P. (2001). Mental Models and Deduction. *Trends in Cognitive Science*, 5(10), 434–442.
- Johnson-Laird, P. (2008). Mental Models and Deductive Reasoning. En J. Adler & L. Rips (Eds.), *Reasoning. Studies of Human Inference and Its Foundations* (pp. 206–222). Cambridge, U.K.: Cambridge University Press.
- Johnson-Laird, P.N. (2012). Inference in Mental Models. En K. Holyoak & R. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 134–154). New York, U.S. A.: Oxford University Press.
- Kahneman, D. & Tversky, A. (1973). *Attention and Effort*. Englewood Cliffs, U.S. A.: Prentice-Hall.
- Kahneman, D. (2012). *Pensar rápido, pensar despacio*. Buenos Aires, Argentina: Editorial Debate.
- Klauer, K. M., Musch, J., & Naumer, B. (2000). On Belief Bias in Syllogistic Reasoning. *Psychological Review*, 107(4), 852–884.
- Khemlani, S. & Johnson-Laird, K. (2017). Illusions in Reasoning. *Minds and Machines*, 27(1), 11–35.
- Krämer, S. (2003). Writing, Notational Iconicity, Calculus: On Writing as a Cultural Technique. *Modern Language Notes*, 118(3), 518–537.
- Oaksford, M. & Charter, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford, U.K.: Oxford University Press.
- Perkins, D. (2002). Standard Logic as a Model of Reasoning: the Empirical Critique. En D. Gabbay, H. J. Ohlbach, R. Johnson, & J. Woods (Eds.), *Handbook of the Logic of Argument and Inference* (pp. 187–224). North Holland, Netherlands: Elsevier.
- Stanovich, K. & West, R. (2000). Individual Differences in Reasoning: Implications for the Rational Debate? *Behavioral and Brain Sciences*, 23(5), 645–665.
- Stanovich, K., West, R., & Toplak, M. (2016). *The Rationality Quotient. Towards a Test of Rational Thinking*. Cambridge (MA), U.S.A.: The MIT Press.
- Stenning, K. & van Lambalgen, M. (2008). *Human Reasoning and Cognitive Science*. Cambridge (MA), U.S.A.: The MIT Press.
- Tarski, A. (1936/2002). On the Concept of Following Logically. *History and Philosophy of Logic*, 23(3), 155–196.
- Toplak, M., West, R., & Stanovich, K. (2014). Assessing Miserly Information Processing: an Expansion of the Cognitive Reflection Test. *Thinking and Reasoning*, 20(2), 147–168.
- Thompson, V. & Evans, J. St B.T. (2012). Belief Bias in Informal Reasoning. *Thinking and Reasoning*, 18(3), 278–230.
- Wason, P. (1966). Reasoning. En B. Foss (Ed.), *New Horizons in Psychology* (pp. 135–151). H. Middlesex, U.K.: Penguin Books.
- Wason, P. (1968). Reasoning about a Rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273–281.
- Woleński, J. (2016). Logic in the Light of Cognitive Science. *Studies in Logic, Grammar and Rhetoric*, 48(61), 87–101.

Notas al final

1 Una línea de investigación alternativa en psicología del razonamiento cuestiona el uso de la lógica de primer orden estándar para analizar el razonamiento humano y propone, en cambio, la teoría de la probabilidad bayesiana para evaluar el razonamiento humano. (Oaksford & Charter, 2007). Este enfoque no será discutido en este trabajo

2 El sesgo focal es independiente de la teoría de modelos mentales. De hecho, podría ser compatible con otras teorías del razonamiento humano como la teoría de la lógica mental (Rips, 1994). Sin embargo, por el enfoque de nuestro trabajo sólo consideraremos el sesgo focal en relación con la construcción de modelos mentales.