

EL DEBATE ACERCA DEL EFECTO FACILITADOR EN PROBLEMAS DE PROBABILIDAD CONDICIONAL: ¿UN CASO DE EXPERIMENTACIÓN CRUCIAL?*

THE DEBATE AROUND THE FACILITATION EFFECT ON CONDITIONAL PROBABILITY PROBLEMS: IS THIS A CASE OF CRUCIAL EXPERIMENTATION?

RODRIGO MORO** Y GUSTAVO A. BODANZA***

*Este trabajo fue financiado con el PICT 693 *Modelos de agentes racionales argumentativos* de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) y por la Secretaría de Ciencia y Técnica de la Universidad Nacional del Sur con el PGI 24/ZI29: *Argumentación y racionalidad* y el PGI 24/ZI38: *Análisis y resolución de debates en psicología cognitiva sobre errores de razonamiento y decisión*.

**Doctor en Filosofía. Profesor Adjunto de la Cátedra Filosofía de la Ciencia del Departamento de Humanidades de la Universidad Nacional del Sur. 12 de Octubre y San Juan, 5to Piso - (8000). Bahía Blanca - Prov. de Buenos Aires. República Argentina. E-Mail: rromo@uns.edu.ar

***Doctor en Filosofía. Profesor Asociado de la Cátedra Lógica del Departamento de Humanidades de la Universidad Nacional del Sur y Miembro de la Carrera del Investigador Científico del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). E-Mail: ccbodanz@criba.edu.ar

RESUMEN

Se ha mostrado en la literatura de *Psicología Cognitiva* que las personas generalmente tienen dificultades para resolver problemas de *probabilidad condicional*. Sin embargo, también se ha mostrado que, bajo ciertas condiciones, las respuestas mejoran de manera significativa. Desde mitad de la década de 1990 hubo un gran debate acerca de cómo dar cuenta de dicho *efecto facilitador*. Se han propuesto dos hipótesis rivales, la *hipótesis de frecuencias naturales* que dice que el efecto facilitador se debe a presentar la información de manera frecuentista, y la *hipótesis de conjuntos anidados* que dice que dicho efecto se debe a la clarificación de las relaciones de conjunto del problema. En este artículo intentaremos clarificar el debate y analizar la evidencia empírica relevante. La pregunta a responder es la siguiente: ¿Se ha producido alguna experimentación crucial en favor de alguna de las dos hipótesis? Nuestra respuesta será negativa, aunque reconociendo que la hipótesis de conjuntos

anidados parece hasta ahora, mejor respaldada que su rival.

Palabras clave: Psicología Cognitiva; Probabilidad condicional; Efecto facilitador; Hipótesis de frecuencias naturales; Hipótesis de conjuntos anidados.

ABSTRACT

In the early '70s, Tversky and Kahneman founded a research program in *Cognitive Psychology* called *Heuristics and Biases*. This program found extensive evidence that shows that people tend to commit reasoning errors when making judgments under uncertainty. A particular case is that people tend to fail when reasoning about *conditional probability* problems, that is, problems that ask for the probability of some event given the fact that another event has occurred (e.g. the probability of rain-

ing given that it is cloudy). But in the mid '90s, Gigerenzer and other evolutionary psychologists came along and gave an important turn to the state of the art. They showed that if the conditional probability problems used in the literature are framed in a different way, people's performance greatly improves. More specifically, if the problems present the information under a specific format called *natural frequency format*, around 50% of participants get the correct answer. Since the mid '90s researchers engage in an important debate on how to account for such a *facilitation effect*. There are two main proposals, one by the Evolutionary Psychology Program and the other by Heuristic and Biases Program. The *natural frequency hypothesis* supported by the *Evolutionary Program* basically says that the natural frequency format is the responsible factor for the improvement in people's performance. The Heuristic and Biases Program, in turn, has proposed the *nested-set hypothesis* to explain the facilitation effect. The basic idea is that natural frequency versions tend to make transparent the relevant subset relations of the problem. When people see clearly the set relations involved in this kind of problems (the argument goes) they tend to use correctly base rates and thus, their performance improves. They point out that, according to this view, the success of the frequency effect does not have to do with natural frequency formats per se. They predict that any format whatsoever that make the relevant set relations clear will show the same effect. The key question is, then, as follows. Is this a case of crucial experimentation in favor of one of our rival hypotheses? In other words, is there an experiment or a series of experiments such that our rival hypotheses predict opposite results, so that we can claim one of them as victorious over the other? The empirical evidence on the matter is mixed. Some studies seem to support the natural frequency hypothesis while others seem to support the nested-set hypothesis. We will then try to clarify this debate by focusing on the diverse strategies and techniques used in the literature to settle the dispute. We will argue that the right strategy to discriminate between both hypotheses is to use genuine probability problems with a clarified set structure and see whether these conditions elicit or not a per-

formance comparable to the natural frequency effect. Within this general strategy, we review the literature and found that there are three techniques, namely, the improved wording technique, the natural chance technique and the graphical representation technique that seem to provoke a performance as good as the one elicited by natural formats, giving, thus, a stronger support for the nested set hypothesis. However, a careful analysis shows that neither the improved wording technique nor the chance technique has provided both consistent and clear results in favor of the nested-set hypothesis. As for the graphical representation technique, the evidence still seems very slim. The improvement in performance was shown in two studies that worked with only one problem each. Furthermore, neither of these problems seems completely adequate. Thus, we do not think the last word about the matter has been said and more empirical work is needed to settle the issue.

Key words: Cognitive Psychology; Conditional probability; Facilitation effect; Heuristic and Biases Program; Evolutionary program; Natural frequency hypothesis; Nested-set hypothesis.

INTRODUCCIÓN

A comienzos de la década de 1970 Amos Tversky y Daniel Kahneman fundaron un programa de investigación en Psicología Cognitiva llamado *Heurísticas y Sesgos* (en adelante, PHS). Este programa ha reportado un gran volumen de evidencia empírica mostrando aparentemente que las personas tienden a cometer errores cuando razonan en situaciones de incertidumbre (ver Gilovich, Griffin & Kahneman, 2002, para una revisión actualizada).

Un caso particular es que la gente usualmente comete errores cuando resuelve problemas de probabilidad condicional. El siguiente es el ejemplo más famoso de este tipo de problemas, denominado el *problema del diagnóstico médico*.

a.- Versión estándar del problema del diagnóstico médico

La probabilidad de que un argentino seleccionado al azar tenga la enfermedad X es 1/1000. Se ha desarrollado un test para detectar dicha enfermedad. Un individuo que no tiene la enfermedad tiene una probabilidad de 50/1000 de que el test le dé positivo. A un individuo que padezca la enfermedad el test definitivamente le dará positivo.

Supongamos que seleccionamos un argentino usando un sorteo de lotería y no sabemos nada sobre los síntomas o signos de esta persona. Asumamos también que a esta persona se le administra el test para la enfermedad y da positivo.

¿Cuál es la probabilidad de que esta persona padezca de hecho la enfermedad? _____ %.

La mayoría de los participantes (médicos incluidos) tienden a responder 95%. La respuesta correcta de acuerdo al cálculo de probabilidades es alrededor de 2%¹. Los estudios muestran que típicamente menos del 20% de los participantes obtiene la respuesta correcta (Casscells, Schoenberger & Grayboys, 1978; Eddy, 1982; Cosmides & Tooby, 1996; Evans, Handley, Perham, Over & Thompson, 2000; Gigerenzer & Hoffrage, 1995; Sloman, Over, Slovak & Stivel (2003). Así, desde el descubrimiento del fenómeno a finales de la década de 1970 parecía claro que la mayoría de la gente (sin el entrenamiento apropiado) era incapaz de solucionar correctamente este tipo de problema.

Pero a mediados de la década de 1990, Gigerenzer y otros psicólogos evolucionistas avanzaron dando un importante giro a la cuestión. Comenzaron notando que un problema como el presentado tiene dos particularidades:

1.- la información es presentada en formato probabilista (por ejemplo, la tasa

base de la enfermedad se informa diciendo que la probabilidad de que un argentino seleccionado al azar tenga la enfermedad X es 1/1000) y

2.- la pregunta formulada es sobre la probabilidad de un evento singular, a saber, la probabilidad de que una persona tenga la enfermedad. Los psicólogos evolucionistas mostraron que si el mismo problema era presentado en términos de frecuencias y no de probabilidades, la respuesta de la gente mejoraba notablemente. Más precisamente, si la información se presentaba en un formato específico llamado de *frecuencias naturales*, alrededor del 50% de los participantes obtenía la respuesta correcta.

b.- Versión frecuentista natural del problema del diagnóstico médico

En el caso del diagnóstico médico, una versión frecuentista natural diría: uno de cada 1.000 argentinos tiene la enfermedad X. Se ha desarrollado un test para detectar dicha enfermedad. Cada vez que el test se aplica a una persona que padece la enfermedad, el test da positivo. Pero a veces el test da positivo cuando se aplica a personas completamente sanas. Específicamente, de cada 999 personas que están completamente sanas, a 50 de ellas el test les da positivo.

Supongamos que seleccionamos 1.000 argentinos al azar. Entre los que les da positivo el test, ¿cuántos padecen la enfermedad de hecho? _____ enfermos de _____ con test positivo.

Nótese que la información no sólo es presentada en formato frecuentista (*Uno de cada 1000 argentinos tiene la enfermedad X*), sino además de frecuencia natural. Tales formatos tienen dos características esencia-

¹ Esto puede verse como una aplicación directa de la regla de Bayes: $P(\text{enfermedad} / \text{positivo}) = \frac{P(\text{enfermedad}) \times P(\text{positivo} | \text{enfermedad})}{P(\text{positivo})} = \frac{0.001 \times 1}{0.001 \times 1 + 0.999 \times 0.05}$.

les: (1) las frecuencias relevantes no están normalizadas (ni en 100 ni en 1.000, por ejemplo) y (2) la información estadística es presentada como una partición o, como a los psicólogos evolucionistas les gusta decir, como si fuera obtenida por muestreo natural, es decir, por actualización de frecuencias de los distintos eventos a medida que la información es obtenida. La gente tiende a responder mucho mejor en estos formatos (ver Gigerenzer & Hoffrage, 1995 para un estudio sistemático). Este fenómeno es ampliamente aceptado en la literatura. La cuestión, que ha sido el principal foco del debate por más de 10 años, es cómo explicar este fenómeno.

DOS EXPLICACIONES RIVALES: LA HIPÓTESIS DE FRECUENCIAS NATURALES VERSUS LA HIPÓTESIS DE CONJUNTOS ANIDADOS

En la literatura hay acuerdo acerca de que la segunda versión es computacionalmente más simple que la primera.

Gigerenzer y Hoffrage (1995) explicitan este punto mostrando que la ecuación necesaria para solucionar la versión frecuentista es más simple (contiene menos elementos y operaciones y utiliza números enteros en lugar de decimales o fracciones) que la ecuación necesaria para solucionar la versión frecuentista. Pero los investigadores del área usualmente van más allá de dicho punto cuando tratan de explicar el fenómeno en cuestión.

Se han planteado dos hipótesis principales acerca de qué factores son los que producen el mencionado efecto facilitador: la *hipótesis de frecuencias naturales* (HFN), sostenida por el Programa Evolucionista (PE) -principalmente por Gigerenzer- que dice que el formato de frecuencia natural es el factor responsable del efecto en cuestión. Este formato requiere que la información y la pregunta del problema estén dados en términos de frecuencias naturales, en lugar de términos de probabilidades. Los defensores del PE incluso especulan con una historia

evolutiva donde nuestros ancestros repetidamente enfrentaron situaciones inciertas y el uso de información frecuentista (por ejemplo: ‘tuvimos éxito 5 de las 20 veces que fuimos a cazar cerca de aquella montaña’) les ayudó a sobrevivir (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Pero más allá del origen del fenómeno, la característica clave es de nuevo, el fraseo en términos de frecuencia natural. A partir de esta hipótesis, el PE predice que la gente se desempeñará mejor si se da la información en tales términos en lugar de darla en términos de probabilidades.

Por otra parte, el Programa de Heurísticas y Sesgos ha propuesto la *hipótesis de conjuntos anidados* (HCA), que dice que el efecto es producido por información que hace claras las relaciones de conjuntos del problema. Las versiones de frecuencia natural tienden a hacer claras dichas relaciones de conjunto y es eso lo que produce el efecto en cuestión. Por ejemplo, en el problema del diagnóstico médico, la versión frecuentista natural hace transparente que el conjunto de gente que obtiene resultados positivos en el test incluye dos subconjuntos, el de gente que padece la enfermedad y el de gente que no la padece (ver Figura 1). Cuando las personas entienden las relaciones de conjunto involucradas, tienden a hacer uso de la información estadística dada y el desempeño mejora. El PHS señala que, de acuerdo con esta hipótesis, la clave del efecto facilitador no radica en el uso del formato de frecuencia natural y predice que cualquier formato que haga transparentes la estructura de conjunto de un problema producirá el mismo efecto facilitador (Giroto & González, 2001; Sloman et al., 2003).

La evidencia empírica en este debate es diversa. Algunos estudios parecen apoyar la hipótesis de frecuencias naturales (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995; Krämer & Gigerenzer, 2005; Zhu & Gigerenzer, 2006), mientras otros parecen apoyar la hipótesis de conjuntos anidados (Evans et al., 2000; Macchi, 2000; Giroto & González, 2001; Sloman et al., 2003; Yamagishi, 2003).

Intentaremos clarificar el debate analizando las estrategias y técnicas usadas por los investigadores del área. Argumentaremos que la evidencia disponible parece favorecer a la hipótesis de conjuntos anidados. Sin embargo, también presentaremos objeciones razonables contra la evidencia provista. Así, concluiremos que la cuestión está todavía abierta y, por lo tanto, no ha habido aún alguna experimentación crucial exitosa.

PRIMERA ESTRATEGIA: USAR FORMATOS FRECUENTISTAS QUE NO TRANSPARENTEN LA ESTRUCTURA DE CONJUNTOS DEL PROBLEMA

Una de las estrategias usadas por los investigadores para decidir entre ambas hipótesis rivales es usar versiones frecuentistas naturales que no clarifiquen la estructura de conjuntos. La idea es que la HFN postularía que, ante un problema presentado en términos frecuentistas pero que no trasluce las relaciones entre las clases involucradas, la gente igualmente tendrá un buen desempeño. Por el contrario, según la HCA, la falta de claridad con respecto a los anidamientos conjuntistas provocarían un mal desempeño. Lamentablemente, esta estrategia no sirve.

La razón es que una de las características principales de las versiones de frecuencias naturales, es que presentan la información de un modo partitivo revelando, en consecuencia, las relaciones conjuntistas. Varios investigadores, sin embargo, han intentado seguir esta estrategia (e.g., Evans et al., 2000). Pero estos intentos han fracasado dado que no se utilizaron frecuencias naturales realmente, sino frecuencias normalizadas (se llevan las proporciones a cifras redondas; por ejemplo, '3 de 8' es normalizado mediante '37 de 100'). Estos estudios muestran que en algunas versiones frecuentistas (no naturales), donde no se revela la estructura de conjunto, las personas se desempeñan muy pobremente. Desafortunadamente, estos resultados no son muy informativos ya que los defensores de la hipótesis de frecuencia natural predicen exactamente el

mismo resultado. Así, esta estrategia no permite distinguir entre nuestras hipótesis rivales.

SEGUNDA ESTRATEGIA: USAR FORMATOS PROBABILÍSTICOS CON UNA CLARA ESTRUCTURA DE CONJUNTO

La estrategia que queda es crear versiones probabilísticas en las que la estructura de conjunto es clarificada y ver si tales versiones producen el efecto facilitador. Esta es la estrategia adecuada porque HFN y HCA predicen resultados opuestos. Si usando estas versiones se produce el efecto facilitador, entonces habría una corroboración de la HCA y se mostraría a la vez que los formatos frecuentistas no son necesarios para producir el efecto. Si no se produce el efecto facilitador, habría una corroboración para la HFN y se mostraría que la clarificación de conjuntos no es suficiente para producir el efecto en cuestión. Poseemos, así, los requisitos necesarios para un experimento crucial.

Sin embargo, aquí surge un importante problema metodológico: cómo distinguir versiones probabilistas de versiones frecuentistas naturales. El problema con esta distinción es que el concepto de probabilidad puede ser interpretado como sugiriendo frecuencias naturales. De hecho, la interpretación frecuentista es una de las interpretaciones más plausible de probabilidad. Además, como se ha notado en la literatura, las mismas expresiones numéricas (porcentajes, fracciones e incluso números enteros) pueden ser legítimamente usados para ambas versiones. En otras palabras, la distinción entre formatos es vaga. No obstante, hay un acuerdo generalizado en la literatura en clasificar ciertos fraseos como típicamente frecuentistas (por ejemplo, 3 de cada 10 casos de A son casos de B) y ciertos fraseos como genuinamente probabilísticos o no-frecuentistas (por ejemplo, la chance o probabilidad de que el evento singular A sea un caso de B es 30%). Desafortunadamente, algunos investigadores no son particularmente cuidadosos al res-

pecto. Algunas versiones pretendidamente probabilísticas, como la de Macchi (2000) se presentan más cercanas a las frecuencias que a las probabilidades (con expresiones como ‘360 de cada 1.000 estudiantes’, o ‘75% de los estudiantes...’, informando un total de 1.000 estudiantes; *ibid.* p. 24, nuestra traducción). Por supuesto, los estudios que usan formatos ambiguos no pueden arrojar resultados claros. El efecto facilitador puede ser producto de la clarificación de la estructura de conjuntos o por la sugerencia de hacer una lectura frecuentista del problema. Esta crítica ha sido repetida por Gigerenzer (2007) y otros autores (Hoffrage, Gigerenzer, Krauss & Martignon, 2002). Lamentablemente, hay varios estudios que son susceptibles de recibir la misma crítica (e.g., Fiedler, Brinkmann, Betsch & Wild, 2000; Sloman et al., 2003).

Teniendo en cuenta esta dificultad metodológica, exploremos las tres técnicas que aplican la estrategia general de usar formatos probabilísticos con una estructura de conjunto clarificada. Estas tres técnicas son en apariencia las más exitosas en mostrar importantes mejoras en el desempeño de la gente, dando así un aparente apoyo a la hipótesis de conjuntos anidados.

Técnica 1: Usar un fraseo mejorado

Sloman y colaboradores (2003) usan distintas versiones del problema del diagnóstico médico, algunas de las cuales tienen un fraseo mejorado que refleja la estructura de conjuntos del problema. Esta técnica puede verse a continuación en la parte italizada.

VERSIÓN MEJORADA DE FORMATO PROBABILÍSTICO CON RELACIONES DE CONJUNTOS ANIDADOS CLARAS

La probabilidad de que un argentino seleccionado al azar tenga la enfermedad X es 1/1000. Se ha desarrollado un test para detectar dicha enfermedad. *Si se usa el test con una persona que padece la enfermedad, el test definitivamente dará positivo. Pero el*

test puede darle positivo a una persona que está completamente sana.

Específicamente, hay una probabilidad de 50/1000 de que una persona perfectamente sana obtenga un resultado positivo en dicho test.

Supongamos que seleccionamos un argentino usando un sorteo de lotería y no sabemos nada sobre los síntomas o signos de esta persona. Asumamos también que a esta persona se le administra el test para la enfermedad y da positivo.

¿Cuál es la probabilidad de que esta persona padezca de hecho la enfermedad?
_____ %.

La idea es enfatizar que el conjunto de personas que obtuvo un test positivo incluye a todos los enfermos, pero también algunos sanos. Si la HCA es correcta, esta versión debería producir significativamente más respuestas correctas que la versión estándar. La HFN, obviamente, predice que no va a haber ningún cambio.

El problema es que Sloman y sus colegas no encontraron resultados consistentes. Ellos usaron varias versiones probabilísticas sin fraseo mejorado y varias con fraseo mejorado (y una versión frecuentista). En algunos casos, encontraron un importante efecto facilitador (de 20% a 48% de respuestas correctas). Sin embargo, usaron varias versiones de cada tipo porque reconocían que algunas de ellas contenían ambigüedades (por ejemplo, versiones probabilísticas que podrían ser interpretadas de manera frecuentística). Ahora bien, hubo una comparación que Sloman y sus colaboradores no hicieron. Es la comparación entre versiones donde las ambigüedades son fuertemente reducidas o eliminadas. Esta es la comparación más relevante.

Las versiones que contienen ambigüedades no son confiables debido a que esas ambigüedades pueden ser fuente de error. Así, la comparación que realmente cuenta es entre las versiones que contienen menos o ninguna ambigüedad. De acuerdo con sus propios criterios, cuando las personas enfrentan la mejor versión probabilística sin claras relaciones de conjunto, el 39% de ellas da la respuesta correcta y cuando enfrentan la mejor versión

probabilística con claras relaciones de conjuntos, el 40% de ellos da la respuesta correcta. Los porcentajes son prácticamente idénticos. Así, para las mejores versiones de cada tipo, la clarificación de las relaciones de conjuntos no parece producir ningún mejoramiento.

Dada la inconsistencia de resultados con esta técnica, no podemos decidir sobre el asunto en juego.

Técnica 2: Usar formato de chance natural

Una de las técnicas más efectivas es la usada por Girotto y González (2001), quienes encontraron una manera muy inteligente de expresar probabilidades de eventos singulares. El siguiente es un ejemplo de su formato de chance natural:

“Los aplicantes para ingresar a una prestigiosa universidad tienen que rendir un examen de ingreso, el cual abarca un test oral y otro escrito. La siguiente es información acerca de los resultados del examen del año pasado:

Un aplicante tenía 5 chances de 100 de ser aceptado. Tres de esas 5 chances estaban asociadas a aprobar el test oral. Siete de las restantes 95 chances de ser rechazado estaban asociadas a aprobar el test oral. Imagina que Juan es uno de los aplicantes para ingresar en dicha universidad. De 100 chances, Juan tiene ___ chances de aprobar el test oral, ___ de las cuales están asociadas con ser aceptado para el ingreso” (p. 272-273, nuestra traducción).

En este formato, la estructura partitiva de la información hace transparentes las relaciones de conjunto del problema. El resultado es que este formato de chance provoca un efecto facilitador similar al de frecuencias naturales. Así, este resultado parece favorecer la hipótesis de conjuntos anidados.

¿Cuál es la respuesta de los defensores de la hipótesis de frecuencias naturales? La objeción principal se basa en la sospecha que las chances no son sino frecuencias dis-

frazadas (Hoffrage et al., 2002). De hecho, la estructura de la versión de chance emula en detalle la estructura de la versión de frecuencias naturales. Además, en la Introducción se menciona que la información es acerca de los resultados del examen del año pasado, lo cual puede inducir también a una lectura frecuentista. Estas observaciones, por supuesto, ponen en duda la interpretación de los resultados como favoreciendo a la HCA. De nuevo, el efecto facilitador podría ser provocado por la clarificación de la estructura de conjuntos o por la pista de leer el problema de manera frecuentista.

De hecho, parece haber cierta evidencia empírica en favor de que se produce, en cierta medida, una lectura frecuentista. La evidencia es aportada por Brase (2008) quien después de presentar a los sujetos el problema del ingreso (p. 285, nuestra traducción) les hace elegir entre las siguientes opciones:

“1.- Pensé en la información como un caso único de aplicación de ingreso con algunas posibilidades de aprobar el test oral y algunas posibilidades de ser aceptado para el ingreso (interpretación probabilística).

2.- Pensé en la información como un gran número de aplicaciones de ingreso, algunas de las cuales aprobaban el test oral y algunas de las cuales eran aceptadas para el ingreso (interpretación frecuentista).

3.- Otra: pensé en la información como ___.”

Brase reporta que un porcentaje importante de participantes (alrededor del 30%) eligió la interpretación frecuentista. Así, aun si la mayoría de los participantes eligió la interpretación probabilística (alrededor del 60%), habría evidencia de que el formato es un tanto ambiguo. Y hay un punto más llamativo: los que eligieron la interpretación frecuentista respondieron significativamente mejor que los que eligieron la interpretación probabilística, por lo que el éxito de este

formato dependería parcialmente de su ambigüedad.

Así, el principal problema del formato de chance es si es un formato genuinamente probabilístico o es un formato frecuentista disfrazado.

Hasta que este punto no sea resuelto, este formato no provee evidencia conclusiva para la hipótesis de conjuntos anidados. Consideremos, entonces, la última técnica que parece discriminar entre nuestras hipótesis rivales.

Técnica 3: Usar representaciones gráficas

Hay una manera alternativa de crear versiones probabilísticas que revelen las relaciones de conjunto del problema: dar una formulación del problema en términos inequívocamente probabilísticos, añadiendo la información acerca de la estructura de conjuntos anidados mediante diagramas de Euler. Esta técnica fue usada por Cosmides y Tooby (1996), Yamagishi (2003) y Sloman y colaboradores (2003). Quizá esta sea la mejor manera de revelar las relaciones de conjunto. Sin embargo, como en las técnicas anteriores, debe aplicarse cuidadosamente para evitar una lectura frecuentista. Este es exactamente el problema con el estudio de Cosmides y Tooby, quienes hacen que los participantes dibujen gráficos en los que aparezca un cuadrado por individuo representado, claramente imponiendo (no solamente sugiriendo) una interpretación frecuentista; Yamagishi (2003) y Sloman y colaboradores (2003), por otra parte, parecen evitar dicho problema. Sloman y sus colaboradores (Experimento 2) usan la siguiente versión del problema del diagnóstico médico:

“Considere una enfermedad que un argentino dado tiene una chance de 1/1000 de ‘pescar’. Un individuo que no padece la enfermedad tiene una chance de 5% de obtener un resultado positivo en el test de la enfermedad. Un individuo que padece la enfermedad definitivamente obtendrá un resultado positivo en

el mismo test. ¿Cuál es la chance de que una persona que obtuvo un resultado positivo padezca de hecho la enfermedad, asumiendo que uno no sabe nada acerca de sus síntomas y signos?” (Sloman et al., 2003, p. 300, nuestra traducción).

Sin dudas, esta versión califica como genuinamente probabilística. Bajo la condición de control, se da a los participantes este problema solo y bajo la condición experimental, se da esta versión acompañada de un gráfico que revela las relaciones de conjunto (ver Figura 1). Sloman y colaboradores reportaron un gran efecto facilitador, yendo de un 20% de respuestas correctas sin diagrama a un 48% con diagrama. Esta mejora fue casi idéntica a la provocada por la versión frecuentista (51%).

Yamagishi (2003) a su vez, reporta un resultado similar. Él usa diferentes variaciones del siguiente problema:

“Una fábrica produce piedras preciosas artificiales. Cada piedra tiene una chance de 1/3 de salir empañada, 1/3 de salir fracturada y 1/3 de salir bien. Una máquina de inspección remueve todas las fracturadas y deja todas las que salieron bien. Sin embargo, la máquina remueve sólo la mitad de las empañadas. ¿Cuál es la chance de que una piedra esté empañada si ha pasado la inspección?” (Yamagishi, 2003, p. 99, nuestra traducción).

En la condición experimental clave, se incluye una representación gráfica como en la Figura 2. Yamagishi reporta una mejora de un promedio de 15% de respuestas correctas sin diagrama y un promedio de 75% con diagrama. En este caso, el efecto del diagrama fue superior al efecto provocado por frecuencias naturales (49%).

Estos dos últimos estudios experimentales proveen la evidencia más clara a favor de la hipótesis de conjuntos anidados y en contra de la de frecuencias naturales. Ninguno de los defensores de esta última hipótesis ha respondido al desafío. Y nótese que propor-

cionar tal respuesta no es nada fácil. Uno puede pensar inicialmente que lo que los defensores de la HFN deben hacer es tratar de mostrar que las representaciones de Sloman y Yamagishi inducen, de alguna manera, a una lectura frecuentista del problema. Pero incluso si este fuera el caso, ninguna de las versiones con gráficos va a ser vista como de frecuencia natural ya que todas las proporciones presentadas están normalizadas. Así, la hipótesis de frecuencias naturales seguiría prediciendo que no habría ninguna mejoría bajo tales condiciones. Por lo tanto, parece muy difícil para los defensores de la HFN incorporar los resultados obtenidos con esta técnica.

Sin embargo, se proporcionan algunas objeciones contra la idea de que existe evidencia conclusiva en favor de la hipótesis de conjuntos anidados (Barbey & Sloman, 2007; Sloman et al., 2003).

ALGUNAS OBJECIONES

La primera objeción tiene que ver con los gráficos usados en Sloman y colaboradores (2003) y Yamagishi (2003) (ver Figuras 1 y 2). Nótese que ambos gráficos muestran algo más que la mera estructura de conjuntos del problema en cuestión. Sucede que también muestran (Yamagishi) o sugieren (Sloman) los tamaños relativos de los conjuntos involucrados. Y esta característica adicional no está presente en la hipótesis de conjuntos anidados tal como se la presenta habitualmente. Uno puede perfectamente presentar la estructura de conjuntos de un problema sin sugerir los tamaños propor-

cionales de las clases en juego (ver Figura 3). Esta característica de revelar las proporciones relativas puede a su vez clarificar el objetivo de la tarea en juego, que es, después de todo, obtener una determinada proporción. Así, puede suceder que lo que provoca el efecto facilitador encontrado no sea la clarificación de la estructura de conjuntos per se, sino la pista acerca de las proporciones relativas. Por supuesto, esto es una especulación y debe ser testado empíricamente para ser tomado seriamente. Aún así, queda abierta la cuestión de que si una representación gráfica no revela las proporciones relativas de sus clases, puede seguir provocando una buena respuesta de parte de los participantes. Si esto no sucediera, tendríamos evidencia negativa en contra de la hipótesis de conjuntos anidados, al menos, tal como es habitualmente presentada.

Una segunda objeción más general puede ser presentada señalando la limitada evidencia en favor de la hipótesis de conjuntos anidados. Como se argumentó anteriormente, las técnicas de fraseo mejorado y la de formato de chance no han podido proporcionar resultados que sean, a la vez, consistentes y claros en favor de la HCA. En cuanto a la técnica de representaciones gráficas, la evidencia aún parece un tanto escasa. El mejoramiento en *performance* fue mostrado en dos estudios que trabajaron con un solo problema cada una. Además, ninguno de estos problemas parece totalmente adecuado. En primer lugar, la versión del problema del diagnóstico médico usada por Sloman es muy problemática como ellos mismos admiten (Sloman et al., 2003, p. 301)². En segundo lugar, el problema de las piedras

² Un primer problema tiene que ver con que la información es dada acerca de una persona 'pescando' una enfermedad sin especificar el período de tiempo que esa enfermedad puede haber sido 'pescada', de tal manera que no es claro que se deba usar la tasa base de la enfermedad para resolver el problema. Un segundo problema radica en que esta versión no aclara que la persona en cuestión es seleccionada al azar. Así, los participantes pueden asumir una probabilidad inicial diferente para la hipótesis de la enfermedad. Por ejemplo, puede suceder que los participantes asuman que el paciente tiene ciertos síntomas que hacen que el médico ordene el test. Si este fuera el caso, uno debería asignar una probabilidad para la hipótesis de la enfermedad mayor a la tasa base (1/1000).

usado por Yamagishi puede ser también cuestionado debido a que las proporciones necesitadas para obtener la respuesta (un tercio, la mitad, etc.) son bastante más simples que las usadas típicamente en la literatura. Así, no queda claro que el fenómeno revelado por la técnica de representaciones gráficas sea robusto, es decir, que persista con otros problemas o con otras versiones de los mismos problemas.

Para resumir, aunque la evidencia provista hasta ahora parece favorecer la hipótesis de conjuntos anidados, es necesario realizar más trabajos empíricos para concluir la disputa en favor de dicha hipótesis.

CONCLUSIÓN

Los problemas de probabilidad condicional son difíciles de resolver. Aun en las condiciones más favorables, la gente debe hacer grandes esfuerzos para hallar el resultado correcto. Sin embargo, se ha mostrado que cuando el problema es presentado bajo un formato de frecuencia natural, el desempeño de la gente mejora. Gigerenzer y otros psicólogos evolucionistas infieren que el uso de formatos de frecuencias naturales es

el factor que causa el efecto de facilitación. Los defensores de la HCA discrepan y postulan una de las características del formato de frecuencias naturales (pero no al formato mismo) como el factor responsable, a saber, la clarificación de las relaciones de conjuntos relevantes del problema. Aquí argüimos que la estrategia correcta para distinguir entre ambas hipótesis es la siguiente: mostrar que una versión genuinamente probabilística del problema con una estructura de conjuntos transparente produce un desempeño similar a las versiones de frecuencias naturales.

Revisamos la literatura y sostuvimos que hay al menos una técnica (la de representaciones gráficas) que parece provocar una mejoría similar a la provocada por formatos de frecuencia natural, dando así apoyo a la hipótesis de conjuntos anidados. Sin embargo, no creemos que la última palabra haya sido dicha. Planteamos algunas objeciones razonables en contra de la posición que sostiene que hay un caso cerrado en favor de la HCA (principalmente, limitaciones de cantidad y calidad de la evidencia encontrada), por lo que resulta necesario diseñar y realizar nuevos experimentos que permitan dirimir la disputa.

FIGURA 1
REPRESENTACIÓN GRÁFICA DEL PROBLEMA DE DIAGNÓSTICO MÉDICO (SLOMAN ET AL., 2003, P. 298)

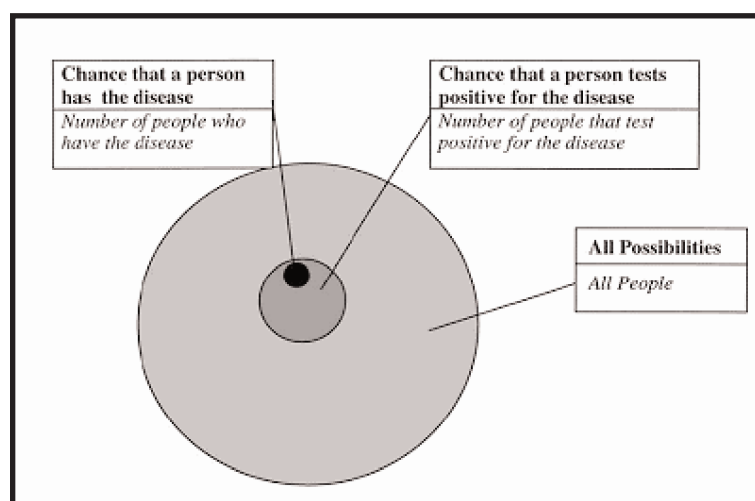


FIGURA 2
REPRESENTACIÓN GRÁFICA DEL PROBLEMA DE LAS PIEDRAS (YAMAGISHI, 2003, P. 98)

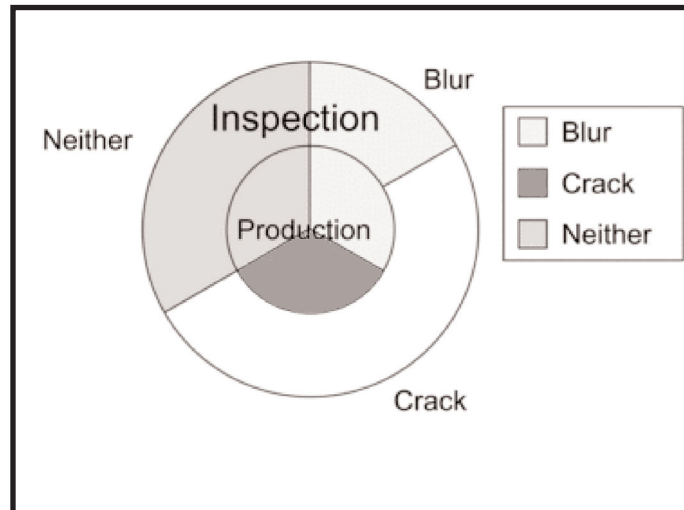
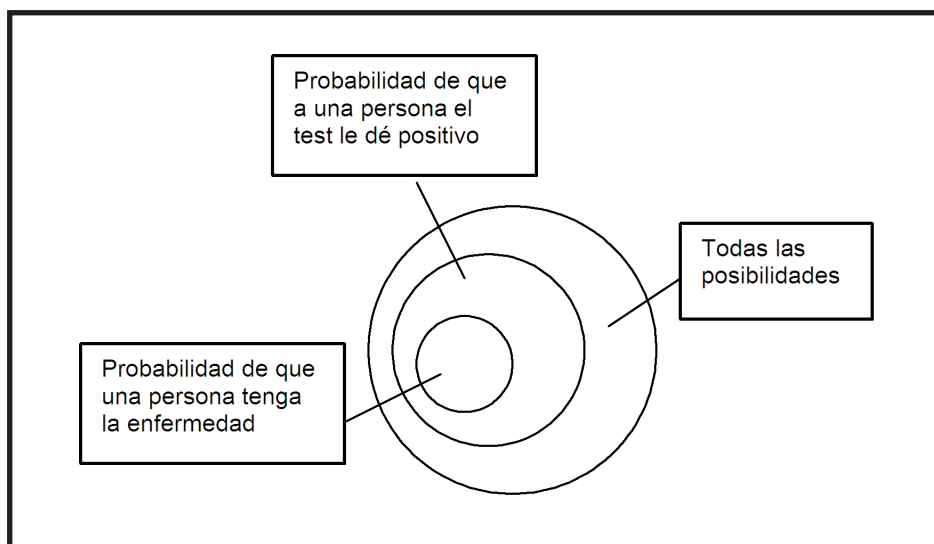


FIGURA 3
REPRESENTACIÓN GRÁFICA DEL PROBLEMA DEL DIAGNÓSTICO MÉDICO QUE NO SUGIERE LOS
TAMAÑOS RELATIVOS DE LAS CLASES INVOLUCRADAS



REFERENCIAS BIBLIOGRÁFICAS

- Barbey, A. & Sloman, S. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, 30(3), 241-254.
- Brase, G.L. (2008). Frequency interpretation of ambiguous statistical information facilitates Bayesian reasoning. *Psychonomic Bulletin & Review*, 15(2), 284-289.
- Casscells, W., Schoenberger, A. & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, 299, 999-1000.
- Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58, 1-73.
- Eddy, D.M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. En D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249-267). Cambridge: Cambridge University Press.
- Evans, J.S., Handley, S.J., Perham, N., Over, D.E. & Thompson, V.A. (2000). Frequency versus probability formats in statistical word problems. *Cognition*, 77, 197-213.
- Fiedler, K., Brinkmann, B., Betsch, T. & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General*, 129(3), 399-418.
- Gigerenzer, G. (2007). The role of representation in Bayesian reasoning: Correcting common misconceptions. *Behavioral and Brain Sciences*, 30(3), 264-267.
- Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102(4), 684-704.
- Gilovich, T., Griffin, D. & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.
- Giroto, V. & González, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. *Cognition*, 78, 247-276.
- Hoffrage, U., Gigerenzer, G., Krauss, S. & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition*, 84, 343-352.
- Krämer, W. & Gigerenzer, G. (2005). How to confuse with statistics or: The use and misuse of conditional probabilities. *Statistical Science*, 20(3), 223-230.
- Macchi, L. (2000). Partitive formulation of information in probabilistic problems: Beyond heuristics and frequency format explanations. *Organizational Behavior and Human Decision Processes*, 82(2), 217-236.
- Mellers, B. & McGraw, P. (1999). How to improve Bayesian Reasoning: Comment on Gigerenzer and Hoffrage (1995). *Psychological Review*, 106, 417-424.
- Sloman, S., Over, D., Slovic, L. & Stivel, J. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, 91(2), 296-309.
- Yamagishi, K. (2003). Facilitating normative judgments of conditional probability: Frequency or nested sets? *Experimental Psychology*, 50(2), 97-106.
- Zhu, L. & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition*, 98, 287-308.

Consejo Nacional de Investigaciones
Científicas y Técnicas (CONICET)
Departamento de Humanidades
Universidad Nacional del Sur (UNS)
Bahía Blanca
Prov. de Buenos Aires - República Argentina

Fecha de recepción: 21 de abril de 2009
Fecha de aceptación: 1 de septiembre de 2009