

ILUSIONES EN EL RAZONAMIENTO PROPOSICIONAL

Carlos Santamaría y Philip N. Johnson-Laird***

*Universidad de La Laguna, **Princeton University

Resumen

Una característica de los modelos mentales como sistema de representación, es que tienden a representar solamente lo que es verdadero. Este supuesto explica que las personas encuentren dificultades para llevar a cabo algunas inferencias que requieren del razonador que tenga en cuenta, no sólo los casos en que el condicional es verdadero, sino aquellos en que es falso. Además de esto, y de forma complementaria, este principio de veracidad predice que las personas llevarán a cabo inferencias ilusorias cuando los casos verdaderos confieran apariencia de validez a ciertas conclusiones falaces. Este caso es particularmente relevante para el actual debate entre modelos mentales y reglas formales de inferencia puesto que las teorías de esta última clase no postulan (ni podrían hacerlo en rigor) reglas paralógicas que den lugar a inferencias ilusorias sistemáticas. Utilizando problemas disyuntivos demostramos que tales ilusiones se producen en el razonamiento proposicional (más del 80% de los sujetos llevan a cabo una inferencia errónea) pero pueden eliminarse sin variar la estructura lógica del problema, haciendo que éste sea resoluble a partir de los casos verdaderos (más del 80% de los sujetos produjeron entonces la inferencia contraria).

Palabras clave: Razonamiento proposicional, modelos mentales, sesgos.

INTRODUCCIÓN

La teoría de modelos mentales (Johnson-Laird y Byrne, 1991) propone que el razonamiento humano se basa en la representación de situaciones que concuerdan con la descripción de las premisas del problema. Por ejemplo, la conjunción: *Hay un cuadrado y hay un círculo*, concuerda con la siguiente situación:



dicho de otro modo, la conjunción es verdadera en el caso de que estén presentes las dos figuras geométricas mencionadas, pero es falsa en cualquier otro caso. Por ejemplo, la frase anterior es falsa en tres tipos de situaciones: cuando haya un cuadrado pero no un círculo, cuando haya un círculo pero no un cuadrado, y cuando no haya ni cuadrado ni círculo. La teoría de los modelos mentales postula que las personas se representan únicamente la situación (o situaciones) en que la proposición es verdadera, lo que supone un sesgo sistemático. En la mayor parte de los casos este sesgo no tiene consecuencias catastróficas sobre la exactitud y corrección del razonamiento puesto que a menudo es posible llegar a la conclusión correcta sin necesidad de tener en cuenta las situaciones falsas. Un claro ejemplo es el *Modus Ponens*:

*Si hay un cuadrado entonces hay un círculo
hay un cuadrado;
por tanto, hay un círculo.*

Existen estructuras en que la inferencia correcta depende de que se tengan en cuenta elementos falsos. Es el caso del *Modus Tollens*:

*Si hay un cuadrado entonces hay un círculo
no hay un círculo;
por tanto, no hay un cuadrado.*

y existe cumplida evidencia de la mayor dificultad de esta inferencia con respecto al *Modus Ponens*. Pero estos datos admiten diversas explicaciones y no resultan de utilidad para discernir, por ejemplo, entre las teorías de reglas formales de inferencia (por ejemplo, Braine y O'Brien, 1991; Rips, 1994) y las de modelos: al fin y al cabo, el que una regla de inferencia esté o no implementada en el sistema cognitivo humano es para los partidarios de las reglas mentales una cuestión puramente empírica, de forma que ellos consideran que el *Modus Ponens* cumple este requisito pero no el *Modus Tollens*, que ha de ejecutarse en varios pasos inferenciales, lo que explicaría su dificultad.

De esta forma, no es posible refutar las teorías de reglas formales demostrando que las personas no ejecutan una inferencia determinada, puesto que siempre puede aducirse que la regla precisa para dicha inferencia no está disponible en el sistema cognitivo. Existe, no obstante, un dato que pondría en cuestión los principios centrales de las teorías de reglas formales de inferencia, y este hallazgo sería encontrar que las personas ejecutan sistemáticamente una inferencia errónea, ya que esto obligaría a las teorías de reglas a postular la presencia de una regla falaz en el sistema cognitivo humano, y la incorporación de tal elemento rompería la consistencia interna del sistema ya que dicha regla entraría en contradicción con el resto de las reglas presentes, dando lugar a un sistema inferencial no computable.

El *principio de veracidad* (las personas tienden a representarse sólo lo que es verdadero) permite predecir que los seres humanos razonarán de forma sistemáticamente falaz cuando un razonamiento basado exclusivamente en las situaciones verdaderas sustente una respuesta errónea, mientras que la obtención de la respuesta correcta sea el resultado de razonar a partir de lo que es falso. Una situación de este tipo fue explorada por Johnson-Laird y Savary (1996) con problemas como este:

Si hay un rey en la mano, entonces hay un as en la mano.
O en caso contrario,
Si no hay un rey en la mano, entonces hay un as en la mano.
Hay un rey en la mano
¿Qué puede deducirse?

Casi todas las personas (incluidos la mayoría de los expertos) concluyen que: *hay un as en la mano*. Sin embargo, la respuesta correcta es justamente la contraria: *no hay un as en la mano*. Pero para llegar a esta conclusión hay que pensar precisamente en los casos falsos. Hay que darse cuenta de que en caso de ser verdadero el primer condicional es falso el segundo, y por lo tanto *no hay un as en la mano*, y en caso de ser verdadero el segundo condicional es falso el primero con lo que de nuevo, *no hay un as en la mano*. Aunque tengamos seguridad de que "*Hay un rey en la mano*" (3ª premisa), no podemos asegurar que haya un as, porque la primera premisa podría ser falsa. La conclusión errónea en estos problemas se nos aparece como verdadera de una forma tan evidente que podemos hablar de una ilusión en el mismo sentido de las ilusiones perceptivas.

Existe, no obstante, un problema en la investigación de Johnson-Laird y Savary que puede comprometer la interpretación de los resultados: podría considerarse que los participantes consideran las premisas inconsistentes y, por tanto, no tienen todas ellas en cuenta. Para descartar esta hipótesis hemos planteado una investigación en la que no utilizamos condicionales sino disyunciones, y donde no existe contradicción interna alguna. Por ejemplo, veamos el siguiente problema:

- (1) Solamente una de estas tres frases es verdadera acerca de Juan:
1. Juan es: físico o ingeniero, o las dos cosas.
 2. Juan es: arquitecto o ingeniero, o las dos cosas.
 3. Juan es: biólogo.
- Juan no es físico y no es arquitecto.

En consecuencia, ¿Es posible que Juan sea ingeniero?

La teoría de los modelos mentales predice que las personas darán mayoritariamente una respuesta afirmativa puesto que pueden producir dos modelos de la primera premisa:

Físico

Ingeniero

Dos de la segunda:

Arquitecto

Ingeniero

Y uno de la tercera:

Biólogo

Puesto que las tres frases están unidas por una disyunción excluyente (enunciada en la expresión "Solamente una de estas tres frases es verdadera") la información se puede integrar en cuatro modelos:

Físico

Ingeniero

Arquitecto

Biólogo

La cuarta premisa niega el primero y el tercero de los modelos:

¬Físico

Ingeniero

¬Arquitecto

Biólogo

Quedando la representación así:

Ingeniero

Biólogo

Con lo que se inferiría que Juan es ingeniero o biólogo: por tanto, es posible que sea ingeniero. Sin embargo, esta respuesta es errónea puesto que si tenemos en cuenta todos los modelos posibles (incluyendo las situaciones falsas) tendríamos que de las dos primeras premisas, al menos una ha de ser forzosamente falsa (podrían serlo las dos si Juan resultase ser biólogo). La primera premisa (una disyunción incluyente) es falsa solamente en el caso de que Juan no sea físico y no sea ingeniero. La segunda premisa es falsa sólo cuando Juan no es arquitecto y no es ingeniero. Por lo tanto, en cualquier

caso, podemos afirmar que Juan no es ingeniero: No es posible que Juan sea ingeniero de acuerdo con las premisas.

Para este problema hemos incluido una condición de control en la que la respuesta correcta es afirmativa y se puede alcanzar con los modelos mentales sin tener en cuenta los casos falsos:

(2) Solamente una de estas tres frases es verdadera acerca de Juan:

1. Juan es: físico o ingeniero, o las dos cosas.
2. Juan es: arquitecto o ingeniero, o las dos cosas.
3. Juan es: biólogo.

Juan no es físico y no es biólogo.

En consecuencia, ¿Es posible que Juan sea arquitecto?

Además de esto, hemos incluido un factor facilitador en el diseño: un antídoto para la ilusión predicho por la teoría de modelos mentales. Con la misma estructura que el problema (1) tenemos:

(3) Juan estaba mirando el periódico en busca de trabajo. Había tres anuncios en la página, pero Juan recortó sólo uno: el único que coincidía con su profesión.

Anuncio 1 solicita: físico o ingeniero, o las dos cosas.

Anuncio 2 solicita: arquitecto o ingeniero, o las dos cosas.

Anuncio 3 solicita: biólogo.

Juan no recortó un anuncio que solicitara un físico y no recortó un anuncio que solicitara un arquitecto.

En consecuencia, ¿Es posible que Juan recortara un anuncio que solicitase un ingeniero?

Consideramos que esta condición establece una separación física (no sólo lógica) entre las distintas proposiciones (1, 2 y 3) que se relacionan mediante la disyunción excluyente. Esto haría que los modelos se mantuvieran separados y las negaciones afectasen a toda la proposición (todo el anuncio) y no a cada elemento:

{ Físico
-

Ingeniero

{ Arquitecto
-

Ingeniero

{ Biólogo

Con lo que se llegaría a la conclusión de que el anuncio era para un biólogo y no podría ser para un ingeniero. También para este problema introdujimos una condición de control.

MÉTODO

Participantes y procedimiento

Participaron en el experimento 32 estudiantes de la Universidad de La Laguna que recibieron un cuadernillo con las instrucciones generales y siete problemas (cada uno en una hoja aparte). Cuatro de los problemas correspondían a las cuatro condiciones descritas anteriormente (Disyunción física, Disyunción física-control, Disyunción lógica, Disyunción lógica-control). Cada problema dentro del cuadernillo se presentó con un contenido distinto: utilizamos dos asignaciones al azar de contenidos obtenidas del conjunto de 16 problemas (cuatro condiciones por cuatro contenidos). Cada participante recibió estos problemas en un orden balanceado de acuerdo con un cuadrado Williams lo que dio lugar a cuatro órdenes fijos en donde todos los problemas aparecieron en cada posición el mismo número de veces en toda la muestra, lo que controla los efectos de arrastre de primer orden.

En una posición fija como segundo, cuarto y sexto problema se incluyeron tres problemas de relleno con una estructura sencilla que llevaba a una respuesta negativa.

Los participantes tenían que responder en cada problema señalando "Sí" o "No". Después de resolver cada problema, los participantes debían señalar su grado de confianza en la respuesta dada en una escala de 5 puntos donde el 1 estaba etiquetado como "Ninguna confianza" y el 5 como "Total confianza".

RESULTADOS

La Tabla 1 presenta los porcentajes de respuestas correctas en cada condición. Se encontró una diferencia fiable entre los problemas experimentales (52%) y de control (100%; Wilcoxon, $z=4,77$; $p<0,0001$, una cola). La diferencia entre las condiciones de disyunción física (92%) y lógica (59%) fue también significativa (Wilcoxon, $z=4,58$; $p<0,0001$, una cola); así como la interacción (Wilcoxon, $z=4,58$; $p<0,0001$, una cola) que se puede explicar por la acusada diferencia entre la disyunción lógica y física en el grupo experimental (Wilcoxon, $z=4,58$; $p<0,0001$, una cola).

El porcentaje de respuestas correctas en los problemas de relleno fue del 95%.

Tabla 1. Porcentajes de conclusiones correctas, y medias del índice de confianza entre paréntesis (N=32).

	Experimental	Control
Dis.-Lógica	19 (4.5)	100 (4.7)
Dis.-Física	84 (4.9)	100 (4.9)

Nota: las respuestas correctas eran: "no" para las condiciones experimentales, y "sí" para las de control.

Para las medidas de confianza, es destacable que la puntuación obtenida fue globalmente muy alta (una media de 4,7 sobre una escala donde el 5 significa "total confianza"). La diferencia entre los problemas experimentales (4,7) y de control (4,8) no fue significativa ($z=1,37$). Sí lo fue sin embargo entre los problemas de disyunción lógica (4,6) y de disyunción física (4,9; Wilcoxon, $z=2,50$; $p<0,02$, dos colas). La interacción no fue significativa ($z=1,10$).

La correlación de Spearman entre la confianza y la corrección de la respuesta no fue significativa ni en la condición experimental de disyunción física ($r=0,16$) ni en la condición experimental de disyunción lógica ($r=0,05$; en las condiciones de control no

se pueden calcular estos coeficientes por ser la corrección de la respuesta una constante).

CONCLUSIÓN

Los resultados de este experimento cuestionan claramente las teorías actuales de reglas formales de inferencia. Por primera vez en la psicología del razonamiento se encuentra que más de un 80% las personas llevan a cabo una inferencia errónea y que, al mismo tiempo, más del 80% realizan la inferencia correcta cuando se realiza una manipulación que no afecta a su estructura lógica. Difícilmente pueden las teorías de reglas formales de inferencia dar cuenta de estos resultados sin recurrir a reglas falaces. Sin embargo, como señala Johnson-Laird (en imprenta), no debemos pensar que estos (o cualesquiera otros) datos refutan globalmente cualquier teoría posible que se basase en reglas formales de inferencia. Ningún resultado experimental puede hacerlo puesto que, en principio, una teoría de reglas formales puede funcionar como un lenguaje de programación en el que se podría implementar cualquier teoría computable (incluida la de modelos mentales)

REFERENCIAS

- Braine, M.D.S., O'Brien, D.P. (1991). A theory of If: A lexical entry, reasoning program and pragmatic principles. *Psychological Review*, 98, 182-203
- Johnson-Laird P.N. (en imprenta). Deductive reasoning. *Annual Review of Psychology*
- Johnson-Laird, P.N., Byrne, R.M.J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum
- Johnson-Laird, P.N., Savary F. (1996). Illusory inferences about probabilities. *Acta Psychologica*, 93, 69-90
- Rips L.J. 1994. *The Psychology of Proof*. Cambridge, MA: MIT Press