

# WHY PSYCHOLOGISTS MUST CHANGE THE WAY THEY ANALYZE THEIR DATA: THE CASE OF PSI: COMMENT ON BEM (2011)

ERIC-JAN WAGENMAKERS, RUUD WETZELS, DENNY BORSBOOM,  
HAN L. J. VAN DER MAAS

2011

Los métodos estadísticos actuales han llevado a la conclusión errónea de que los poderes psíquicos son reales. Los resultados de Ben (2011), que indican que existe algo tal como la precognición con evidencia significativa, más que hacernos reconsiderar lo que sabemos de parapsicología, deberían hacernos reconsiderar lo que sabemos de estadística. Los métodos estadísticos actuales son demasiado maleables, simples, y ofrecen demasiadas oportunidades para que los investigadores se engañen a sí mismos.

Sus principales problemas son: confusión entre investigación exploratoria y confirmatoria, desconocimiento de que  $P(\text{Datos} \mid \text{Hipótesis})$  no es lo mismo que  $P(\text{Hipótesis} \mid \text{Datos})$  (i.e., la falacia del condicional traspuesto), y la aplicación de un test que sobreenfatiza la evidencia contra la hipótesis nula. Cuando se aplica una prueba T Bayesiana, la evidencia favorece levemente a  $\mathcal{H}_0$  en algunas ocasiones, y a  $\mathcal{H}_1$  en otras, pero siempre en un grado “anecdótico”, es decir, que apenas vale la pena mencionarse.

## PROBLEMA 1. EXPLORACIÓN EN LUGAR DE CONFIRMACIÓN

Los psicólogos rara vez hacen investigación puramente confirmatoria. Bem mismo sugiere que, dada la desconexión de muchos investigadores con su proceso de recogida de datos, al menos se deberían tomar la molestia de explorarlos a fondo, más allá de los análisis estadísticos clásicos. Sin embargo, Wagenmakers sugiere que, si bien se puede hacer eso, es necesario hacer una distinción explícita entre los resultados dados por los procedimientos convencionales y aquellos que resultan de salir a pescar a ver qué se encuentra en los datos. Cuando un investigador reporta los resultados de esas salidas de pesca como si fueran confirmatorios, está siendo un puerco, porque oculta el hecho de que usó los mismos datos dos veces: primero para descubrir una nueva hipótesis y después para probarla.

En lugar de ser un puerco, se debe usar un procedimiento de dos pasos: primero, en ausencia de una hipótesis, se exploran los datos hasta poder proponer una. Después, con datos distintos, se pone a prueba. Los análisis estadísticos de los estudios exploratorios deben ser más conservadores.

Los estudios de Bem fueron al menos en parte exploratorios. Probaron una variedad

de diferentes clases de imágenes como estímulo hasta dar con la conclusión de que las imágenes eróticas producían la precognición. Sin embargo, de haber encontrado resultados distintos, otra suposición habrían hecho, cuando lo correcto sería realizar un segundo experimento, ya con la hipótesis en mente, para confirmar los hallazgos.

Esto es síntoma de un problema más profundo aún: no hay manera de saber cuántos otros factores se tomaron en consideración antes de llegar a las variables “correctas”. Quién sabe cuántas otras hipótesis fueron probadas y descartadas.

No se realizaron experimentos estrictamente confirmatorios, y por lo tanto, su valor de  $p$  debe ser ajustado hacia arriba.

### PROBLEMA 2. LA FALACIA DEL CONDICIONAL TRASPUESTO

En esta falacia, la probabilidad de los datos dada la hipótesis (i.e.,  $P(\mathcal{D} | \mathcal{H})$ ) es confundida con la probabilidad de la hipótesis dados los datos (i.e.,  $P(\mathcal{H} | \mathcal{D})$ ).

De esta distinción surge el principio de Laplace de “declaraciones extraordinarias requieren evidencia extraordinaria”. Así, la probabilidad previa adjunta a una hipótesis afecta la fuerza de la evidencia requerida para que un observador racional cambie de parecer.

Dada la evidencia en contra de los fenómenos psíquicos, el prior que representa nuestra creencia en ellos debería estar muy cerca de cero, por ejemplo, en 0.00000000000000000001. Ahora, aún encontrando evidencia fuerte a favor de la existencia de fenómenos psíquicos, por ejemplo, evidencia que fuese 19 veces más verosímil bajo  $\mathcal{H}_0$  que bajo  $\mathcal{H}_1$ , la actualización del prior estaría dada por

$$\begin{aligned} p(\mathcal{H} | \mathcal{D}) &= \frac{p(\mathcal{D} | \mathcal{H}_1)p(\mathcal{H}_1)}{p(\mathcal{D} | \mathcal{H}_0)p(\mathcal{H}_0) + p(\mathcal{D} | \mathcal{H}_1)p(\mathcal{H}_1)} \\ &= \frac{.95 \times 10^{-20}}{.05(1 - 10^{-20}) + .95 \times 10^{-20}} \\ &= .00000000000000000019 \end{aligned}$$

Aunque la creencia posterior es 19 veces mayor que la previa, aun tenemos casi certeza sobre la inexistencia de los fenómenos psíquicos. Se requeriría de evidencia mucho más fuerte para dejar atrás el escepticismo. Es decir, declaraciones extraordinarias requieren evidencia extraordinaria.

Así, aun los ocho experimentos de Bem que favorecen la precognición no son suficientes para convencer a nadie, solo en parte por ser a la vez exploratorios y confirmatorios y por tener evidencia débil.

### PROBLEMA 3. LOS VALORES DE $p$ SOBREENFATIZAN LA EVIDENCIA CONTRA LA HIPÓTESIS NULA

Un valor de  $p = 0.001$ , que indica una baja probabilidad de encontrar los datos dada una cierta  $\mathcal{H}_0$ , NO es indicador de que deba aceptarse  $\mathcal{H}_1$ . Los datos podrían ser igual de inverosímiles bajo la hipótesis alternativa. El problema subyacente es que la probabilidad es relativa, y no es muy útil conocer la probabilidad de ciertos datos dada una única hipótesis. Así, para evaluar la fuerza de cierta evidencia a favor de la precognición, la

hipótesis nula debe compararse con una hipótesis alternativa específica y no verse en el vacío. Los Bayesianos lo logran usando pruebas de hipótesis que computan una tasa de verosimilitud ponderada.

La meta en las pruebas de hipótesis Bayesianas es cuantificar el cambio en la probabilidad posterior que es traído por los datos. Para una elección entre  $\mathcal{H}_0$  y  $\mathcal{H}_1$ , tenemos:

$$\frac{p(H_0 | D)}{p(H_1 | D)} = \frac{p(H_0)}{p(H_1)} \times \frac{p(D | H_0)}{p(D | H_1)}$$

Es decir,

Probabilidades del modelo posterior = Probabilidades del modelo prior  $\times$  Factor de Bayes

El Factor de Bayes a menudo es interpretado como el peso de la evidencia dada por los datos. Cuando el factor de Bayes de  $H_0$  sobre  $H_1$  es de 2, significa que los datos son dos veces más probables bajo  $H_0$  que bajo  $H_1$ . Al aplicarse una prueba T Bayesiana, que compara la verosimilitud de  $H_0$  contra  $H_1$  y no toma en cuenta las probabilidades prior, se encontró que la evidencia a favor de la existencia de los poderes psíquicos es, en el mejor de los casos, anecdótica.

#### **GUÍAS PARA LA INVESTIGACIÓN CONFIRMATORIA**

La investigación exploratoria es insuficiente como evidencia para convencer de la plausibilidad de fenómenos que de antemano parecen inverosímiles.

La característica básica de los estudios confirmatorios es que todas las elecciones que podrían afectar los resultados se hacen antes de que los datos se observen. Se sugieren las siguientes guías:

1. Las “expediciones de pesca” se deben evitar seleccionando participantes e items antes del estudio.
2. Los datos deben transformarse solo si esto se ha decidido de antemano. Esto significa que un estudio confirmatorio fallido no se vuelve exploratorio, ni uno exploratorio exitoso se vuelve confirmatorio.
3. En situaciones simples, el análisis apropiado debe elegirse antes de la recolección de los datos.
4. Se deberían reportar múltiples análisis estadísticos, especialmente si sus resultados discrepan.

Estas medidas suelen ser suficientes. Sin embargo, para fenómenos especialmente increíbles se debe ir más allá:

5. Los investigadores deben hacer sus materiales (estímulos, programas y datos) públicamente disponibles.
6. Los investigadores deberían colaborar con escépticos declarados y calificados, preferiblemente más de una vez.

#### **CONCLUSIONES**

Aunque estos experimentos no son prueba de precognición, son evidencia de que el estándar para la evidencia en el ámbito académico podría estar demasiado abajo. Es fácil culpar a Bem por su metodología, pero él se apegó a todas las reglas impuestas implícitamente para la publicación académica. De hecho presentó más estudios de los que se le hubiesen podido exigir.

Más que un ataque a la investigación de fenómenos inusuales, este estudio señala una falla crítica dentro de la forma en que los psicólogos experimentales diseñan, analizan y reportan sus estudios. Es preocupante que quizá muchos fenómenos aceptados tengan evidencia similar, basada en investigación sesgada y exploratoria.