

# ACTAS

## III Jornadas de Investigación en Humanidades



Bahía Blanca  
1 al 3 de octubre de 2009

## Midiendo el entendimiento de las relaciones de conjunto: algunos resultados empíricos

Rodrigo Moro  
UNS / CONICET  
[rmoro@uns.edu.ar](mailto:rmoro@uns.edu.ar)

Gustavo Bodanza  
UNS / CONICET  
[ccbodanz@criba.edu.ar](mailto:ccbodanz@criba.edu.ar)

Esteban Freidin  
CONICET  
[efreidin@criba.edu.ar](mailto:efreidin@criba.edu.ar)

### 1. Introducción

A partir de los estudios de Tversky & Kahnemann en los '70 se ha mostrado que la gente tiende a razonar equívocamente con problemas de probabilidades condicionales como el siguiente:

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 de 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 a 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%. Los estudios muestran que típicamente menos del 20% de los participantes obtiene la respuesta correcta (Casscells *et al.*, 1978; Gigerenzer & Hoffrage, 1995; Cosmides and Tooby, 1996; Sloman *et al.*, 2003). Así, desde el descubrimiento del fenómeno a finales de los '70s 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 los '90 Gerd Gigerenzer y otros psicólogos evolucionistas avanzaron dando un importante giro a la cuestión. Comenzaron notando que un problema como el presentado arriba 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 obtenían la respuesta correcta. En el caso del diagnóstico médico, una versión frecuentista natural diría:

Uno de cada 1000 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 1000 argentinos al azar. Entre los que les da positivo el test, ¿cuántos padecen la enfermedad de hecho?  
\_\_\_\_\_ enfermos de \_\_\_\_\_ con test positivo.

Se han planteado dos hipótesis principales para dar cuenta de dicho efecto facilitador: la que llamaremos *Hipótesis de Conjuntos Anidados (HCA)*, propuesta desde el *Heuristics and Biases Programm* (ver Gilovich *et al.*, 2002), que dice que el efecto es producido por información que hace claras las relaciones de conjuntos del problema, y la que llamaremos *Hipótesis de las Frecuencias Naturales (HFN)*, sostenida desde los '90 por la psicología evolucionista –principalmente por Gerd Gigerenzer– que dice que el efecto es producido por una presentación del problema en términos de frecuencias naturales. La evidencia empírica en este debate es diversa. Algunos estudios parecen soportar la *HFN* (Cosmides and Tooby, 1996; Gigerenzer and Hoffrage, 1995; Krämer and Gigerenzer, 2005) mientras otros parecen soportar la hipótesis de conjuntos anidados (Giroto and González, 2001; Yamagishi, 2003; Sloman *et al.*, 2003). Recientemente, Barbey y Sloman (2007) revisan la literatura y concluyen que la evidencia empírica favorece a la *HCA*. El problema es que, hasta ahora, el factor de entendimiento de relaciones de conjunto no ha sido empíricamente medido y pensamos que sería importante hacerlo para evaluar con propiedad la hipótesis en cuestión. A continuación reportamos dos experimentos donde medimos el factor de entendimiento de las relaciones de conjunto y evaluamos su impacto en la resolución de problemas de probabilidad condicional. Argumentaremos que nuestros resultados empíricos arrojan evidencia negativa en contra de la hipótesis de conjuntos anidados.

## 2. Experimento 1: midiendo el entendimiento de las relaciones de conjunto

El objetivo de nuestro primer experimento fue medir el entendimiento de las relaciones de conjunto para luego poder comparar esos resultados con el desempeño de otros participantes en problemas de probabilidad condicional como el del diagnóstico médico.

Nuestros participantes fueron 92 estudiantes de la carrera de Contador Público de la UNS quienes se hallaban cursando la materia “Introducción al Estudio de las Ciencias Sociales” en el primer cuatrimestre de 2009. El experimento fue realizado en dos sesiones de alrededor 45 participantes cada una. No se impusieron restricciones de tiempo. Cada sesión duró unos 30 minutos.

El material constaba de 2 problemas (el del diagnóstico médico y otro) donde a los participantes se les daba la información de dichos problemas pero, en lugar de pedir juicios probabilísticos, se les hacía una serie de preguntas acerca de las relaciones de

conjunto clave. Había 4 condiciones básicas obtenibles de la combinación de formato (probabilista/frecuentista) y presencia o ausencia de un gráfico que revelaba las relaciones de conjuntos. Cada participante respondió 3 tipos de preguntas sobre las relaciones de conjunto. El primer tipo contenía preguntas de verdadero-falso de proposiciones condicionales (ej. “Si el test da positivo, entonces la persona seguro está enferma”). El segundo tipo de preguntas contenía representaciones gráficas (véase figura 1) donde los participantes debían señalar la o las representaciones que se correspondieran con la información dada. El tercer tipo de preguntas era de verdadero-falso sobre proposiciones conjuntistas (ej. “El grupo de los positivos está completamente incluido en el de los enfermos”). Cada participante obtuvo un puntaje correspondiente al porcentaje de respuestas correctas. Los resultados del test pueden apreciarse en la figura 2. Como puede apreciarse, en general, la gente entendió bastante bien las relaciones de conjuntos de ambos problemas (un promedio general de 86% de respuestas correctas). Nótese que en las versiones clave, es decir, en las probabilísticas, el gráfico parece haber ayudado, aunque sólo significativamente en el problema del diagnóstico médico. La clave, sin embargo, está en chequear si estos resultados se corresponden con el desempeño en los problemas de probabilidad condicional. Eso lo analizamos en Experimento 2.

### **3. Experimento 2: estudiando habilidad en problemas de probabilidad condicional**

Este experimento tiene la misma estructura que el anterior pero, en lugar de las preguntas sobre relaciones de conjunto, requerimos los típicos juicios de probabilidad condicional. Nuestros participantes fueron 80 estudiantes de la Carrera de Contador Público de la UNS de la misma materia pero de distinta comisión que los del experimento previo. El experimento también tomó lugar en el primer cuatrimestre de 2009. No se impusieron límites de tiempo y la sesión duró unos 40 minutos. Los resultados pueden verse en la tabla 1. Nótese la gran discrepancia con los resultados del experimento anterior. Si entender las relaciones de conjunto fuera la clave para resolver los problemas de probabilidad condicional –como postula la hipótesis de conjuntos anidados– uno esperaría que la gran mayoría participantes resolvieran los problemas correctamente. Sin embargo, notamos que este no fue el caso. Si bien las versiones frecuentistas tendieron a generar respuestas adecuadas, este no fue el caso con las versiones probabilísticas. Otro punto crucial es que la presencia del gráfico en el problema del diagnóstico produjo un mejoramiento significativo en el entendimiento de las relaciones de conjunto. Sin embargo, este mejoramiento no se vio reflejado en la resolución del problema probabilístico. Consideramos que esta falta de paralelismo entre los resultados de los dos experimentos provee evidencia negativa para la hipótesis de conjuntos anidados.

### **4. Conclusión**

Los problemas de las probabilidades condicionales son difíciles de resolver. Aún en las condiciones de más ayuda, 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.

Gerd Gigerenzer y otros psicólogos evolucionistas infieren de esto 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 a 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 para el problema. Reportamos una serie de experimentos que testean esta última hipótesis. En primer lugar, investigamos empíricamente el entendimiento de las relaciones de conjunto. En segundo lugar, analizamos si este factor podía dar cuenta del desempeño en problemas de probabilidad condicional. El resultado fue negativo, por lo que consideramos la hipótesis de conjuntos anidados debe o bien ser reformulada o bien ser desechada como explicación del efecto facilitador en cuestión. Por supuesto, esta conclusión no provee una prueba definitiva para la hipótesis de frecuencias naturales ya que las versiones de frecuencias naturales son más simples de resolver que cualquier versión probabilística y el factor simplicidad podría ser el responsable del efecto facilitador. Por lo tanto, más investigación empírica es requerida para dirimir la disputa.

### **Bibliografía**

- Barbey, A. y Sloman, S. (2007), "Base-rate respect: From ecological rationality to dual processes", en *Behavioral and Brain Sciences*, vol. 30, n° 3, pp. 241-254.
- Casscells, W., Schoenberger, A., y Grayboys, T. (1978), "Interpretation by physicians of clinical laboratory results", en *New England Journal of Medicine*, vol. 299, pp. 999-1000.
- Cosmides, L., y Tooby, J. (1996), "Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty", en *Cognition*, vol. 58, pp. 1-73.
- Gigerenzer, G., y Hoffrage, U. (1995), "How to improve Bayesian reasoning without instruction: Frequency formats", en *Psychological Review*, vol. 102, n° 4, pp. 684-704.
- Gilovich, T., Griffin, D., y Kahneman, D. (2002), *Heuristics and biases: The psychology of intuitive judgment*, Cambridge, Cambridge University Press.
- Giroto, V., y Gonzalez, M. (2001), "Solving probabilistic and statistical problems: A matter of information structure and question form", en *Cognition*, vol. 78, pp. 247-276.
- Hoffrage, U., Gigerenzer, G., Krauss, S., y Martignon, L. (2002), "Representation facilitates reasoning: what natural frequencies are and what they are not", en *Cognition*, vol. 84, pp. 343-352.
- Sloman, S., Over, D., Slovak, L., y Stivel, J. (2003), "Frequency illusions and other fallacies", en *Organizational Behavior and Human Decision Processes*, vol. 91, n° 2, pp. 296-309.
- Yamagishi, K. (2003), "Facilitating normative judgments of conditional probability: Frequency or nested sets?", en *Experimental Psychology*, vol. 50, n° 2, pp. 97-106.

Figura 1: Ejercicio involucrando representaciones gráficas donde los participantes debían encerrar en un círculo el o los gráfico/s que representen correctamente las relaciones entre salud, enfermedad y resultados del test:

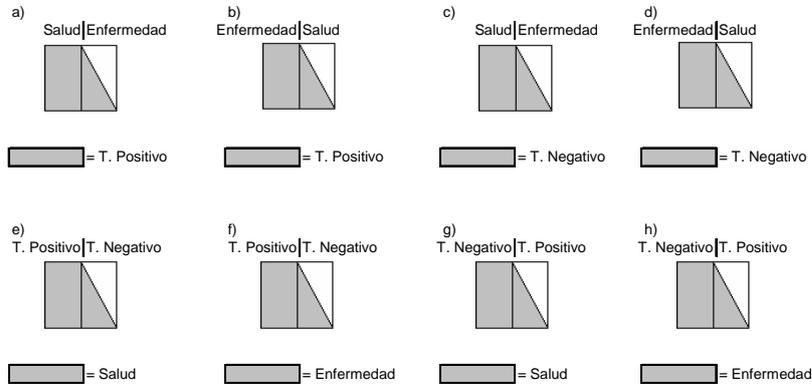


Figura 2: Resultados del Experimento 1

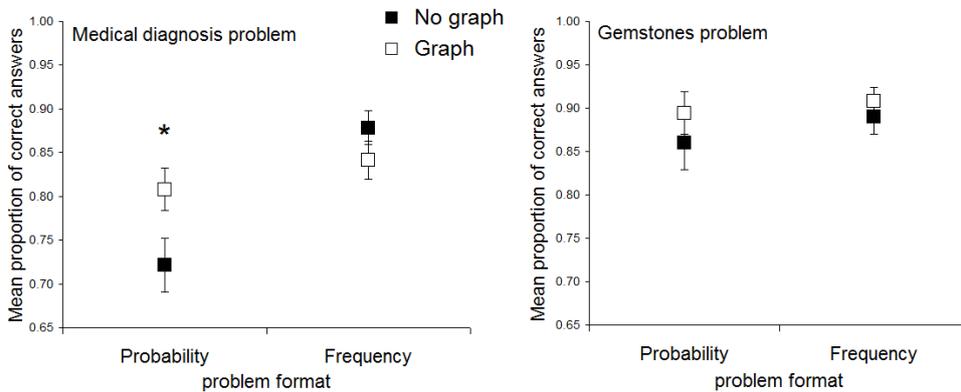


Tabla 1: Resultados del experimento 2

	a) Medical diagnosis problem		b) Gemstones problem	
	No Graphic	Graphic	No Graphic	Graphic
Probability	5.3 (1/19)	0.0 (0/20)	4.3 (1/23)	4.8 (1/21)
Frequency	33.0 (7/21)	50.0 (10/20)	88.9 (16/18)	72.2 (13/18)