

Una teoría naturalista y dinámica de las jerarquías de evidencia

XXIV Jornadas Rolando Chuaqui 2024

π *Jornadas*
ROLANDO CHUAQUI KETTLUN
filosofía y ciencias

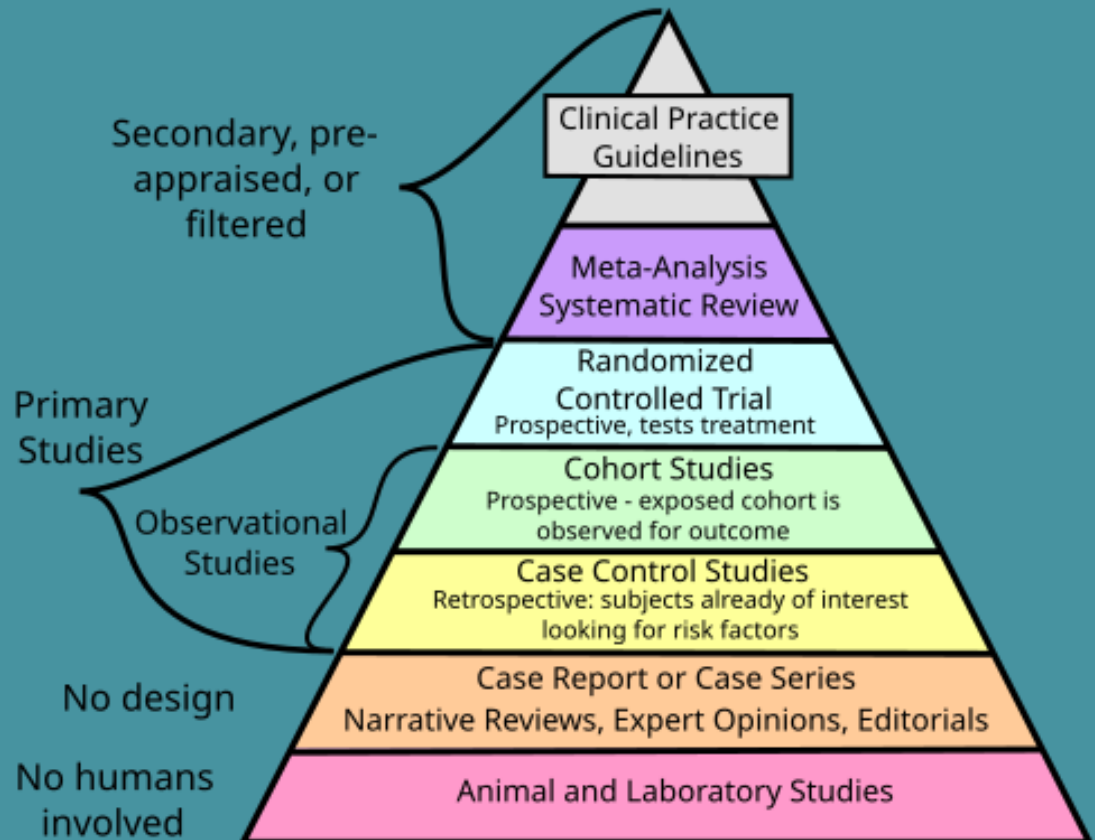


Lucas Miranda Baños – Faro UDD



1. Introducción

Jerarquías de evidencia en EBM y EBP:



Tres literaturas con un tópicocomún

1. Filosofía de la experimentación científica (Hacking, 1983)
2. Jerarquías de evidencia en EBM y EBP (Worrall, 2007; Cartwright, 2007, 2010)
3. El problema de la diversidad de la evidencia causal (Russo & Williamson, 2007)

El tópicocomún:

La evaluación epistémica de distintos tipos de evidencia, en especial la experimental vs la observacional

Tres problemas en torno a este tópico:

1. Comparabilidad de los *comparanda*
2. Clasificación de los experimentos naturales
3. La falacia del nirvana evidencial

La solución será una teoría de las jerarquías de evidencia:

- a) Naturalista: le quita la potestad al filósofo de juzgar las jerarquías de evidencia y se la reconoce exclusivamente en las instituciones científicas.
- b) Social: reserva para el filósofo *qua* epistemólogo social una facultad crítica de las instituciones epistémicas que obstaculizan o favorecen el proceso de jerarquización.
- c) Dinámica: reserva para el filósofo un rol especulativo acerca de la tendencia que tendrá el proceso jerarquizador de evidencia.

Contenidos

1. Introducción
2. Las literaturas y los problemas
3. Datos, fenómenos y error
4. Experimentos ideales y reales
5. El emprendedor metodológico
6. El rol del filósofo

2. Las literaturas y los problemas

1. Filosofía de la experimentación científica

- Remontándose a Hacking (1983) destaca el rol de la experimentación como objeto de análisis filosófico, en contraste al énfasis puesto en las teorías científicas por el empirismo lógico y sus críticos.
- Esta literatura suele plantear alguna forma de control y manipulación como elemento distintivo de la evidencia experimental (Boyd, 2023; Boyd & Matthiessen, 2024; Brandon, 1994; Hacking, 1983; Perović, 2021; Radder, 2009) y en base a ello discute los méritos relativos de los distintos tipos de evidencia.
- Así por ejemplo, en el campo de lo que Currie & Levy (2019) han llamado la “visión tradicional” están todos quienes le asignan a la capacidad manipulativa de los experimentos el sustento de su superioridad epistémica respecto de la observación (Currie & Levy, 2019; Hacking, 1983, 1989; Okasha, 2011; Zwier, 2013).
- Mientras que otros cuestionan esta superioridad identificando rasgos distintos a la manipulación como los relevantes para la superioridad epistémica (Boyd, 2023; Boyd & Matthiessen, 2024; Malik, 2017).
-

Comparabilidad de los *comparanda*

- *EH*: La evidencia de tipo e_1 se debe comparar con la evidencia de tipo e_2 respecto a una misma hipótesis h .
- En las ciencias empíricas la investigación testea a menudo lo que Hacking (1983) llamó *topical hypothesis*.
- Las hipótesis tópicas contrastadas por distintos estudios con diversos métodos evidenciales suelen variar, por lo que no se cumple *EH*.
- *EC*: La calidad del tipo de evidencia e_1 que se compara con el tipo e_2 debe ser la misma pero a menudo no lo es.

Clasificación de los experimentos naturales:

Si lo que define a un experimento es el control y la manipulación humana con fines epistémicos, entonces los experimentos naturales son experimentos en un sentido subrogado o metafórico.

Sin embargo, se utiliza esta metáfora porque tiene alguna característica que la manipulación experimental suele producir de manera más regular.

Por otro lado, hay manipulaciones de un fenómeno que no constituyen un experimento.

Sugerencia: lo característico de los experimentos no es la manipulación, sino que otras propiedades del fenómeno que se producen de manera más segura y regular a través de la manipulación.

2 y 3. Jerarquías y diversidad de evidencia

Literatura que discute la validez de las jerarquías de evidencia que se han generado en el movimiento de Medicina Basada en la Evidencia (EBM) y las Políticas Públicas Basadas en la Evidencia (EBP).

Las influentes críticas de Worrall (2007), Cartwright (2007, 2010), Cartwright & Munro, (2010) y Deaton & Cartwright (2018) han generado una serie de respuestas que defienden o critican estas jerarquías.

Un tercer conjunto de literatura discute acerca de qué hacer ante la diversidad de evidencia en medicina y ciencias sociales y que hace falta para aseverar la existencia de una relación causal.

La tesis Russo-Williamson (Russo & Williamson, 2007) defiende un pluralismo evidencial consistente en exigir la necesaria colaboración de evidencia estadística y mecanística para establecer relaciones causales.

La falacia del nirvana evidencial

La crítica de estos autores a las jerarquías de evidencia y en específico a la evidencia experimental en las ciencias sociales y de la salud se suele hacer comparando la evidencia producida con un ideal metodológico que no tiene ejemplificación en la práctica (Backmann, 2017)

Autores clásicos de la filosofía de la ciencia como Popper, Kuhn y Lakatos han señalado convincentemente que una teoría, paradigma o programa de investigación no es reemplazado cuando experimenta falsaciones o anomalías, sino solo cuando existe una alternativa que pueda reemplazarlo.

Análogamente, se puede decir que una metodología de investigación es superada o al menos subjerarquizada, solo cuando existen alternativas capaces de reemplazarlas.

Los filósofos críticos de las jerarquías de evidencia no siempre acompañan sus críticas con reemplazos.



SCIENTIFIC CONTRIBUTION

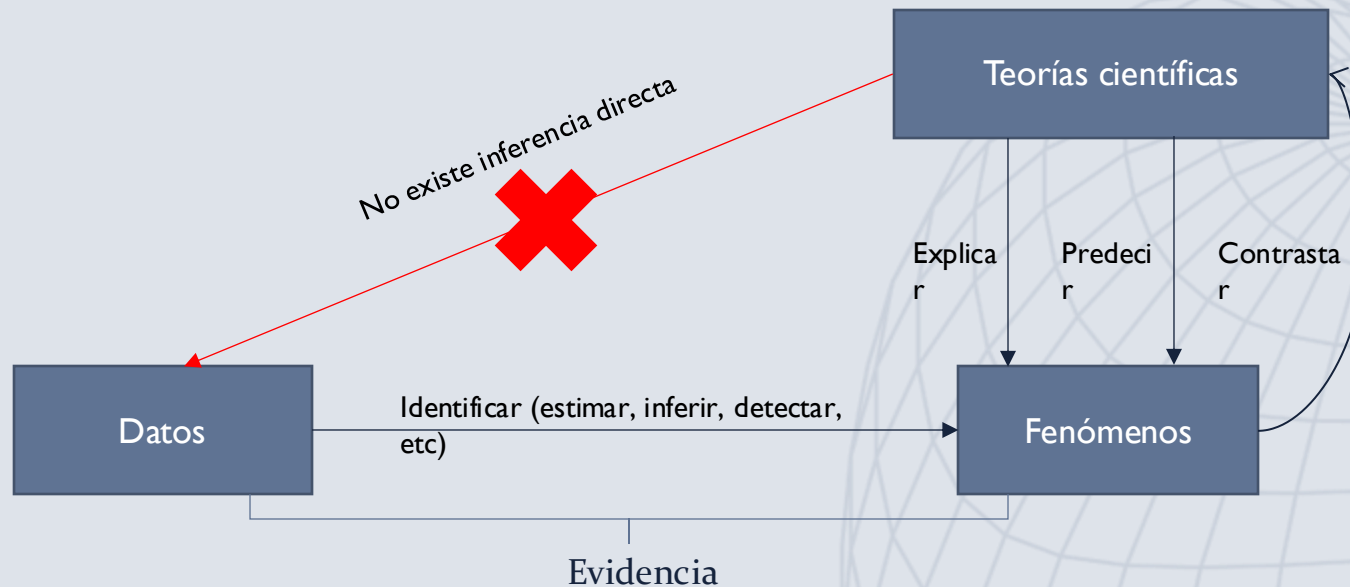
What's in a gold standard? In defence of randomised controlled trials

Marius Backmann¹ 

- Crítica de Worrall (2002): hay demasiados posibles confounders y no re-corremos la prueba para saber si nuestro grupo es atípico. Al haber tantos posibles confounders es probable que la agrupación esté skewed respecto al menos alguno de ellos.
- Crítica de Cartwright (2007): los RCTs pretenden establecer relaciones causales a través de relaciones deductivas basadas en una teoría probabilística de la causalidad.
- En ambos casos se evalúan los RCTs en contra de un ideal que se les achaca y no contra la performance de los estudios observacionales.

3. Datos, fenómenos y error

Bogen y woodward (1988)



Inductivismo eliminativista (Guala, 2005; Reiss, 2007)

¿Qué hace falta para que e sea una buena pieza de evidencia para concluir algo acerca de h ?

Descartar posibles fuentes de error (e.g., hipótesis alternativas) en nuestra inferencia de h a partir de e .

¿Cómo descartarlos?

Identificando los posibles errores que se pueden cometer en una situación empírica dada el identificando métodos para eliminarlos o disminuirlos.

Errores tipificables y particulares

Las teorías metodológicas (estadística, econometría, diseño experimental, etc.,) contienen en gran medida una teoría de los errores idiosincráticos que se pueden cometer en determinada situación empírica infiriendo h a partir de e .

Contienen además métodos para domesticar estos errores (Miranda, 2022).

Siempre hay fuentes de error particulares que solo la experiencia y habilidad del investigador puede abordar.

Errores idiosincráticos de estudios observacionales evitados con RCTs

- *Sesgo de auto-selección*
- *Sesgo de omisión de variables*
- *Efecto placebo (cuando el RCT es ciego)*
- *Sesgo de adherencia*



4. Experimentos ideales y reales

Experimentos ideales y distinción con la evidencia observacional

Haavelmo (1944) y Angrist & Pischke (2009) plantean que ante un determinado problema de investigación el primer paso es plantear el experimento ideal que permitiría responderlo.

Los fenómenos tienen modos de ocurrencia habituales que producen errores idiosincráticos en la inferencia a partir de datos.

El experimento ideal es una forma idealizada de ocurrencia del fenómeno bajo la cual se elimina la fuente de los errores idiosincráticos y se considera una eliminación exhaustiva de las fuentes de error.

Propuesta de distinción: (i) evidencia observacional: la que se obtiene a partir de la forma habitual de ocurrencia del fenómeno; (ii) evidencia experimental: la que se obtiene a partir de una forma de ocurrencia del fenómeno que elimina los errores idiosincráticos, *según una teoría del error típico*.

Solución preliminar de dos problemas

Se soluciona el problema de la clasificación de los experimentos naturales: la distinción ya no requiere 'manipulación'.

Se soluciona *aparentemente* el problema de la comparabilidad de los *comparanda*: se está comparando directamente dos tipos de evidencia, producidas por dos formas de ocurrencia del mismo fenómeno.

Sin embargo...

Problema de circularidad: se está inscribiendo la superioridad epistémica de la evidencia experimental en la propia distinción con la observacional, lo cual resuelve el problema de las jerarquías de evidencia de manera trivial.

Solución: la posible inferioridad de la evidencia experimental respecto de la observacional proviene de dos fuentes: (i) los experimentos reales se alejan de los ideales en formas que introducen errores más perniciosos que algunos errores idiosincráticos; (ii) la teoría del error sobre la cual se basa la distinción tiene problemas.

La crítica a determinada jerarquía de la evidencia puede basarse en cualquiera de estos dos puntos.

5. El emprendedor metodológico

No me gusta la expresión



¿Quién es el agente de esta crítica?

La crítica en la ciencia requiere no solo señalar las falencias de lo que se critica, sino que exhibir la superioridad de una alternativa.

El filósofo *qua* filósofo puede hacer lo primero pero no lo segundo, por lo cual su crítica es fútil o cae en la Falacia del Nirvana.

Sin embargo, los propios científicos tienen incentivos para, dado el estado del arte, hallar fuentes de error en la investigación presente y generar alternativas reconocidas como superiores por los pares.

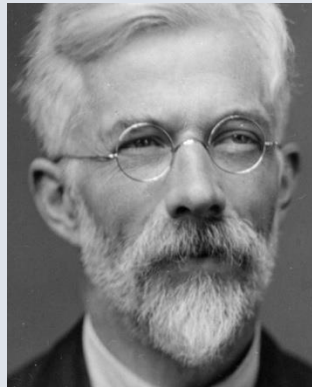
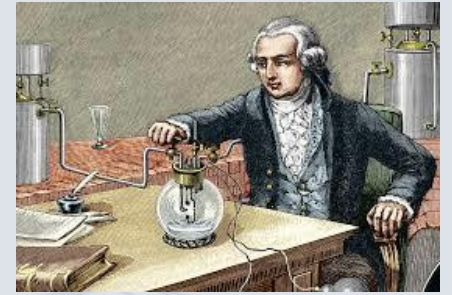
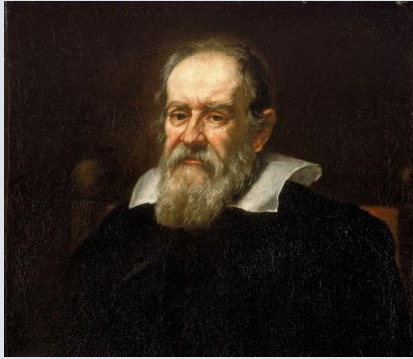
La alerta a la oportunidad de hacer esto se asemeja a la teoría del *entrepreneur* de algunas corrientes económicas.

El *entrepreneur* metodológico

Hay alerta a oportunidades metodológicas que mejoran la práctica dentro de los estándares metodológicos presentes y otras que los cambian.

Es cierto que existen científicos, revistas e instituciones que no innovan metodológicamente y se acoplan de manera pasiva a los estándares de su disciplina.

Pero hay división del trabajo en la ciencia y algunos investigadores e instituciones son innovadores (Kitcher 1990; Strevens 2003; Zollman 2010; Hong & Page 2004)



6. El rol del filósofo

Epistemología social

El filósofo aún podría reservarse un rol crítico pero indirecto de las jerarquías de la evidencia.

Evaluar las instituciones epistémicas, incentivos de los investigadores, su interacción, etc.

Preguntas: ¿Lo anterior obstaculiza o facilita el trabajo del emprendedor metodológico? ¿Cómo se puede complementar la labor del EM?

Especulación dinámica

Preguntas: ¿Las características de una disciplina permiten anticipar conjeturalmente el tipo de evidencia que prevalecerá?

Puede que en un momento un tipo de evidencia no sea la mejor pero que se apueste a que prevalecerá

Por ejemplo, por agotamiento del uso de evidencia de menor calidad

Bibliografía

- Angrist, J., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Backmann, M. (2017). What's in a gold standard? In defence of randomised controlled trials. *Medicine, Health Care, and Philosophy*, 20(4), 513–523.
- Beach, D. (2021). Evidential pluralism and evidence of mechanisms in the social sciences. *Synthese*, 199(3), 8899–8919.
- Bogen, J., & Woodward, J. (1988). Saving the Phenomena. *The Philosophical Review*, 97(3), 303–352.
- Boyd, N. M. (2023). Laboratory Astrophysics: Lessons for Epistemology of Astrophysics. In N. Mills Boyd, S. De Baerdemaeker, K. Heng, & V. Matarese (Eds.), *Philosophy of Astrophysics: Stars, Simulations, and the Struggle to Determine What is Out There* (pp. 13–32). Springer International Publishing.
- Boyd, N. M., & Matthiessen, D. (2024). Observations, Experiments, and Arguments for Epistemic Superiority in Scientific Methodology. *Philosophy of Science*, 91(1), 111–131.
- Brandon, R. N. (1994). Theory and experiment in evolutionary biology. *Synthese*, 99(1), 59–73.
- Cartwright, N. (2007). Are RCTs the gold standard? *BioSocieties*, 2(1), 11–20.
- Cartwright, N. (2010). What are randomised controlled trials good for? *Philosophical Studies*, 147(1), 59–70.
- Cartwright, N. (2021). Rigour versus the need for evidential diversity. *Synthese*, 199(5), 13095–13119.
- Cartwright, N., & Munro, E. (2010). The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16(2), 260–266.
- Currie, A., & Levy, A. (2019). Why experiments matter*. *Inquiry*, 62(9–10), 1066–1090.
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science and Medicine*, 210, 2–21.
- Ghiara, V. (2022). Taking the Russo-Williamson thesis seriously in the social sciences. *Synthese*, 200(6), 481.
- Guala, F. (2005). *The Methodology of Experimental Economics*. Cambridge University Press.
- Haavelmo, T. (1944). The Probability Approach in Econometrics. *Econometrica*, 12, iii–115.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge university press.

- Hacking, I. (1989). Extragalactic Reality: The Case of Gravitational Lensing. *Philosophy of Science*, 56(4), 555–581.
- Kuhn, T. S. (1997). *The structure of scientific revolutions* (Vol. 962). University of Chicago press Chicago.
- Laudan, L. (1984). *Science and Values: The Aims of Science and Their Role in Scientific Debate*. University of California Press.
- Laudan, L. (1987). Progress or Rationality? The Prospects for Normative Naturalism. *American Philosophical Quarterly*, 24(1), 19–31.
- Laudan, L. (1990). Normative Naturalism. *Philosophy of Science*, 57(1), 44–59.
- Malik, S. (2017). Observation Versus Experiment: An Adequate Framework for Analysing Scientific Experimentation? *Journal for General Philosophy of Science*, 48(1), 71–95.
- Maziarz, M. (2021). Resolving empirical controversies with mechanistic evidence. *Synthese*, 199(3), 9957–9978.
- Miranda Baños, L. (2022) Modelos estadísticos y explicación del error inferencial. (2022). *Culturas Científicas*, 3(2), 151-167.
- Morgan, M. S. (2013). Nature’s Experiments and Natural Experiments in the Social Sciences. *Philosophy of the Social Sciences*, 43(3), 341–357.
- Okasha, S. (2011). Experiment, observation and the confirmation of laws. *Analysis*, 71(2), 222–232.
- Perović, S. (2021). Observation, Experiment, and Scientific Practice. *International Studies in the Philosophy of Science*, 34(1), 1–20.
- Radder, H. (2009). The philosophy of scientific experimentation: A review. *Automated Experimentation*, 1(1), 2.
- Reiss, J. (2007). *Error in Economics: Towards a More Evidence-Based Methodology*. Routledge.
- Runhardt, R. W. (2022). Limits to evidential pluralism: Multi-method large-N qualitative analysis and the primacy of mechanistic studies. *Synthese*, 200(2), 171.
- Russo, F., & Williamson, J. (2007). Interpreting Causality in the Health Sciences. *International Studies in the Philosophy of Science*, 21(2), 157–170.
- Shan, Y., & Williamson, J. (2022). Evidential monism, evidential pluralism, or evidential contextualism? An introduction to evidential diversity in the social sciences. *Synthese*, 200(4), 321.
- Taylor, S. D. (2021). Causation and cognition: An epistemic approach. *Synthese*, 199(3), 9133–9160.
- Woodward, J. F. (2011). Data and phenomena: A restatement and defense. *Synthese*, 182(1), 165–179.
- Worrall, J. (2007). Why There’s No Cause to Randomize. *The British Journal for the Philosophy of Science*, 58(3), 451–488.
- Zwier, K. R. (2013). An Epistemology of Causal Inference From Experiment. *Philosophy of Science*, 80(5), 660–671.