

METACIENCIA OPERATORIA:

Auditoría, cierre y decisión en la ciencia contemporánea

Ensayo académico

Autor: Mtro. Julio David Rojas

www.ateliervelazquez.com

12/27/2025

1. Índice

1. Índice
2. Resumen y palabras clave
3. Introducción
4. Preguntas detonantes
5. Contexto y antecedentes históricos
6. ¿Qué es la metaciencia operatoria?
7. Núcleo técnico: cinco pilares de la metaciencia operatoria
8. Exponentes y corrientes principales
9. El Materialismo Filosófico como marco de cierre
10. Mini-caso aplicado: Economía conductual
11. Patologías: no-cierre y sobrecierre
12. Conclusión
13. Glosario
14. Referencias

www.ateliervelazquez.com

2. Resumen y palabras clave

Resumen

La metaciencia contemporánea ha documentado extensamente las fallas sistémicas de la producción científica: tasas de replicación decepcionantes, sesgos de publicación, flexibilidad analítica y estructuras de incentivos perversas. Sin embargo, este diagnóstico permanece incompleto si no se traduce en criterios operatorios de decisión. El presente ensayo propone la noción de metaciencia operatoria como aquel ejercicio de segundo orden que no se limita a describir la ciencia, sino que establece umbrales para la acción bajo incertidumbre epistémica. Tomando como marco el Materialismo Filosófico de Gustavo Bueno, se argumenta que la noción de cierre categorial ofrece un criterio gnoseológico robusto para determinar cuándo un cuerpo de evidencia alcanza la suficiencia operatoria necesaria para justificar intervenciones prácticas. Se analizan los cinco pilares de la metaciencia operatoria —reproducibilidad, preregistro, potencia estadística, trazabilidad y gobernanza— y se ilustra su aplicación mediante un mini-caso en economía conductual. Finalmente, se identifican las patologías simétricas del no-cierre y el sobrecierre, advirtiendo que la ausencia de criterios formales conduce tanto a la parálisis decisional como al dogmatismo institucional.

Palabras clave

Metaciencia operatoria, reproducibilidad, cierre categorial, Materialismo Filosófico, preregistro, potencia estadística, gobernanza científica.

3. Introducción

La metaciencia, entendida como el estudio sistemático de la ciencia mediante métodos científicos, ha experimentado un auge notable en las últimas dos décadas. Desde el artículo seminal de Ioannidis (2005), que argumentaba por qué la mayoría de los hallazgos publicados son probablemente falsos, hasta los esfuerzos coordinados de replicación masiva de la Open Science Collaboration (2015), el campo ha acumulado un diagnóstico devastador sobre el estado de la producción científica contemporánea. La psicología social replica apenas un tercio de sus efectos clásicos; la investigación preclínica en cáncer fracasa en más del 80% de los intentos de reproducción; las ciencias económicas exhiben tamaños de efecto inflados que se desvanecen bajo escrutinio riguroso. Este paisaje de ruinas epistémicas ha generado una industria diagnóstica floreciente: artículos sobre sesgo de publicación, flexibilidad analítica, *p-hacking*, *HARKing* y prácticas cuestionables de investigación llenan las páginas de las revistas metodológicas.

Sin embargo, existe una asimetría fundamental entre la capacidad diagnóstica de la metaciencia y su potencia prescriptiva. Sabemos que la ciencia está enferma, pero carecemos de criterios claros para determinar cuándo un cuerpo de evidencia es suficiente para actuar. Esta laguna no es meramente técnica; es profundamente filosófica. La metaciencia descriptiva nos dice qué está mal; la metaciencia operatoria debe decirnos qué hacer con esa información. Mientras la primera acumula metaanálisis y estadísticas preocupantes, la segunda debe enfrentar la pregunta ineludible: ¿cuándo está justificado actuar sobre una base de evidencia imperfecta, y cuándo la prudencia exige suspender el juicio?

La tesis central de este ensayo es que **la metaciencia contemporánea adolece de un déficit operatorio que solo puede subsanarse mediante la introducción de criterios filosóficos robustos de cierre y suficiencia**. La noción de cierre categorial, desarrollada por Gustavo Bueno en el marco del Materialismo Filosófico, ofrece precisamente el aparato conceptual necesario para distinguir entre estados epistémicos que justifican la acción y estados que exigen cautela o pausa. No se trata de importar acríticamente una filosofía de la ciencia como autoridad retórica, sino de mostrar que sus categorías resuelven problemas operatorios que la metaciencia empírica no puede abordar con sus propios recursos.

La distinción entre metaciencia descriptiva y metaciencia operatoria puede ilustrarse mediante una analogía médica. Un diagnóstico exhaustivo que enumera síntomas, identifica patógenos y describe mecanismos fisiopatológicos es condición necesaria pero no suficiente

para el tratamiento. El clínico debe además decidir cuándo intervenir, con qué intensidad, y bajo qué criterios considerar el tratamiento exitoso o fallido. Esta dimensión decisional es precisamente lo que falta en gran parte de la literatura metacientífica contemporánea. Se nos dice que los efectos de priming son frágiles, pero no cuándo debería un psicólogo clínico abandonar intervenciones basadas en ellos. Se documenta que los ensayos preclínicos en oncología fracasan sistemáticamente, pero no cuándo debería una agencia reguladora modificar sus estándares de evidencia.

El presente ensayo se propone, por tanto, desarrollar la noción de metaciencia operatoria mediante tres movimientos argumentativos. Primero, se contextualiza históricamente el surgimiento de la metaciencia como respuesta a crisis epistémicas acumuladas. Segundo, se presentan los cinco pilares técnicos que constituyen el núcleo de la práctica metacientífica: reproducibilidad, preregistro, potencia estadística, trazabilidad y gobernanza. Tercero, se introduce el Materialismo Filosófico de Gustavo Bueno como marco filosófico que permite articular estos elementos técnicos en un criterio coherente de cierre operatorio. Finalmente, se aplican estas categorías a un caso concreto —la economía conductual— para mostrar cómo la metaciencia operatoria puede guiar decisiones prácticas sobre la suficiencia de la evidencia.

4. Preguntas detonantes

El desarrollo de una metaciencia genuinamente operatoria exige confrontar un conjunto de preguntas que la literatura actual evade o responde solo parcialmente. Estas preguntas no son meramente retóricas; constituyen los ejes problemáticos que estructuran el presente ensayo y que cualquier propuesta de metaciencia operatoria debe abordar con rigor.

1. **¿Cuándo es racionalmente justificado actuar sobre evidencia científica imperfecta?** Toda evidencia es, en algún sentido, imperfecta; la pregunta operatoria es cuándo las imperfecciones son tolerables y cuándo exigen suspensión del juicio o investigación adicional.
2. **¿Qué constituye un cierre suficiente en una disciplina científica?** Si el cierre absoluto es inalcanzable, ¿existen grados de cierre que permitan distinguir entre campos maduros y campos en formación? ¿Cómo se operacionaliza esta distinción?
3. **¿Cuáles son los límites epistémicos del consenso científico?** El consenso puede ser producto de convergencia genuina de evidencias independientes o de dinámicas sociológicas (sesgos de publicación, conformismo institucional, efectos de cascada informativa). ¿Cómo distinguir entre ambos?
4. **¿Qué papel desempeña la reproducibilidad en la validación del conocimiento?** La irreproducibilidad puede indicar error, fraude o variabilidad legítima. ¿Bajo qué condiciones la falla de replicación constituye evidencia contra una teoría y cuándo refleja meramente limitaciones metodológicas?
5. **¿Qué relación existe entre verdad científica y utilidad práctica?** Una teoría puede ser aproximadamente verdadera pero prácticamente inútil, o aproximadamente falsa pero instrumentalmente valiosa. ¿Debe la metaciencia operatoria priorizar la verdad o la eficacia?
6. **¿Cuáles son las patologías específicas del exceso y del defecto de cierre?** El no-cierre conduce a fragmentación y parálisis; el sobrecierre genera dogmatismo y resistencia al cambio. ¿Existen indicadores que permitan diagnosticar cada patología antes de que se vuelva irreversible?
7. **¿Es posible formalizar computacionalmente los criterios de cierre operatorio?** Si la metaciencia operatoria ha de ser verdaderamente operatoria, sus criterios deben ser

aplicables de manera sistemática y, eventualmente, algorítmica. ¿Qué condiciones debe cumplir tal formalización?

5. Contexto y antecedentes históricos

5.1. De Merton a la meta-research contemporánea

La reflexión sistemática sobre la ciencia como objeto de estudio tiene raíces que se remontan a la sociología de la ciencia de Robert K. Merton en la década de 1940. Merton identificó las normas que supuestamente regulan la práctica científica —universalismo, comunismo epistémico, desinterés y escepticismo organizado— como el ethos distintivo de la empresa científica. Sin embargo, la sociología mertoniana operaba en un nivel normativo-descriptivo que asumía, más que demostraba, la efectividad de estas normas para producir conocimiento fiable.

La transición hacia una metaciencia propiamente empírica comenzó con los estudios de citación de Derek de Solla Price en la década de 1960 y se consolidó con el desarrollo de la bibliometría como disciplina cuantitativa. No obstante, estos enfoques permanecían fundamentalmente descriptivos: medían patrones de producción y difusión del conocimiento sin pronunciarse sobre su calidad epistémica. La pregunta de si la ciencia cumplía sus promesas de generar conocimiento fiable quedaba fuera del alcance de estas herramientas.

El giro hacia la metaciencia crítica se produjo gradualmente durante las décadas de 1990 y 2000, impulsado por acumulación de evidencia sobre fallas sistemáticas. Los trabajos de Douglas Altman sobre deficiencias en el reporte de ensayos clínicos, las investigaciones de John Ioannidis sobre sesgos en la literatura médica, y los análisis de estadísticos como Andrew Gelman sobre las limitaciones del paradigma de pruebas de hipótesis nulas convergieron en un diagnóstico cada vez más sombrío. La publicación de «*Why Most Published Research Findings Are False*» (Ioannidis, 2005) cristalizó este malestar en una tesis provocadora que desafiaba los fundamentos mismos de la práctica investigativa contemporánea.

5.2. La crisis de reproducibilidad como evento histórico

Lo que hoy denominamos *crisis de reproducibilidad* no fue un descubrimiento repentino sino la culminación de décadas de preocupaciones acumuladas. Sin embargo, ciertos eventos funcionaron como catalizadores que transformaron inquietudes dispersas en una crisis

institucional reconocida. El artículo de Begley y Ellis (2012) reportando que científicos de Amgen solo pudieron reproducir 6 de 53 estudios preclínicos «landmark» en oncología fue particularmente impactante por provenir de la industria farmacéutica, cuyo interés comercial en replicar hallazgos prometedores era evidente. Si ni siquiera los resultados más celebrados resistían escrutinio riguroso, ¿qué podía esperarse del corpus general?

El proyecto de replicación masiva de la Open Science Collaboration (2015) proporcionó la confirmación sistemática que la evidencia anecdótica había anticipado. De 100 estudios publicados en revistas prestigiosas de psicología, solo el 36% replicaron con criterios estrictos. Los tamaños de efecto promedio cayeron a la mitad. Algunos efectos emblemáticos —como el priming social de John Bargh— se desvanecieron completamente bajo protocolos más rigurosos. La replicación en economía experimental por Camerer et al. (2018) reveló patrones similares: aproximadamente el 60% de los estudios replicaron, pero con efectos sustancialmente menores.

Estos proyectos transformaron la crisis de reproducibilidad de un fenómeno discutido principalmente en círculos metodológicos a un problema reconocido institucionalmente. Agencias de financiación comenzaron a exigir planes de manejo de datos; revistas implementaron políticas de preregistro; se crearon iniciativas como el Center for Open Science dedicadas a reformar las prácticas investigativas. Sin embargo, como argumentaremos, estas respuestas institucionales permanecen fundamentalmente reactivas y descriptivas, sin articular criterios claros de cuándo la evidencia acumulada justifica intervenciones prácticas.

5.3. Manifestaciones disciplinarias

La crisis de reproducibilidad se manifestó de manera diferenciada según las disciplinas. En **psicología social**, los efectos más vulnerables fueron aquellos basados en manipulaciones sutiles de contexto (priming) y los que postulaban influencias inconscientes sobre la conducta. La literatura sobre *ego depletion*, por ejemplo, pasó de ser un paradigma dominante a un caso de estudio sobre inflación de efectos. Los estudios de poder estadístico de Button et al. (2013) demostraron que la neurociencia típicamente operaba con potencias tan bajas que los hallazgos positivos eran más probablemente falsos positivos que descubrimientos genuinos.

En **biomedicina**, la crisis adquirió dimensiones particularmente graves dado que las decisiones clínicas dependen de la solidez de la evidencia preclínica. El fracaso documentado por Begley y Ellis no era meramente académico: cada estudio fallido representaba

potencialmente años de desarrollo farmacéutico desperdiciado y, en última instancia, pacientes que no recibieron tratamientos efectivos porque los recursos se canalizaron hacia líneas de investigación espurias.

En las **ciencias sociales cuantitativas**, incluyendo economía y ciencia política, el problema adoptó formas distintivas relacionadas con la especificación de modelos y el *p-hacking*. El trabajo de Simmons, Nelson y Simonsohn (2011) demostró experimentalmente cómo decisiones analíticas aparentemente inocuas —exclusión de valores atípicos, selección de covariables, definición de variables dependientes— podían inflar dramáticamente las tasas de falsos positivos, haciendo que incluso datos aleatorios produjeran resultados «significativos».

www.ateliervelazquez.com

6. ¿Qué es la metaciencia operatoria?

6.1. Definición formal

Metaciencia operatoria: Disciplina de segundo orden que estudia sistemáticamente la producción, validación y aplicación del conocimiento científico con el propósito explícito de establecer *criterios de decisión bajo incertidumbre epistémica*. A diferencia de la metaciencia descriptiva, que se limita a documentar características de la práctica investigativa, la metaciencia operatoria determina umbrales de suficiencia evidencial que justifican la acción práctica, la cautela provisional o la suspensión del juicio.

Esta definición requiere desempaque. La expresión «disciplina de segundo orden» indica que la metaciencia toma como objeto de estudio no fenómenos naturales o sociales directamente, sino las prácticas mediante las cuales otras disciplinas estudian tales fenómenos. El énfasis en «criterios de decisión» distingue la metaciencia operatoria de aproximaciones puramente contemplativas o historiográficas. La referencia a «incertidumbre epistémica» reconoce que toda decisión basada en evidencia científica opera bajo condiciones de conocimiento incompleto; la pregunta no es si existe incertidumbre sino cómo manejarla racionalmente.

6.2. Diferenciación de disciplinas afines

La metaciencia operatoria debe distinguirse de varias disciplinas con las que comparte territorio pero de las cuales difiere en objetivos y métodos. Respecto a la **filosofía de la ciencia**, la diferencia radica en que esta última tradicionalmente ha operado en un registro normativo-abstracto, estableciendo criterios de demarcación, racionalidad y progreso que frecuentemente carecen de aplicabilidad directa. La metaciencia operatoria, sin renunciar a la reflexión filosófica, exige que sus categorías sean traducibles en procedimientos concretos de evaluación.

Respecto a la **sociología de la ciencia**, la distinción es más sutil. Ambas disciplinas estudian empíricamente la práctica científica, pero la sociología tiende hacia el relativismo epistémico, tratando la «verdad» como construcción social sin privilegio ontológico. La metaciencia operatoria, en cambio, asume que existe una diferencia real entre conocimiento fiable y artefactos metodológicos, y que esta diferencia puede diagnosticarse mediante criterios intersubjetivamente válidos.

Respecto a la **bibliometría**, la diferencia es de alcance. La bibliometría mide patrones de publicación, citación e impacto sin pronunciarse directamente sobre calidad epistémica. Un artículo altamente citado puede ser influyente sin ser verdadero; un campo bibliométricamente productivo puede estar generando principalmente ruido. La metaciencia operatoria utiliza datos bibliométricos como insumos pero los integra en evaluaciones más comprehensivas de solidez evidencial.

6.3. El problema operatorio fundamental

El problema operatorio fundamental puede formularse así: dado un cuerpo de evidencia científica sobre una cuestión prácticamente relevante, ¿bajo qué condiciones está justificado actuar como si esa evidencia fuera aproximadamente verdadera? Esta pregunta es irreducible a cualquiera de las disciplinas mencionadas porque involucra simultáneamente componentes empíricos (¿qué dice realmente la evidencia?), metodológicos (¿cómo fue producida?), sociológicos (¿qué incentivos moldearon su producción?) y filosóficos (¿qué criterios de suficiencia aplicamos?).

Munafò et al. (2017) propusieron un «manifiesto por la ciencia reproducible» que identificaba amenazas a la investigación científica y soluciones institucionales. Sin embargo, este manifiesto permanecía fundamentalmente descriptivo: enumeraba problemas y proponía remedios sin articular criterios de cuándo considerar que un cuerpo de evidencia había alcanzado robustez suficiente. La metaciencia operatoria requiere precisamente lo que el manifiesto omite: umbrales de cierre que permitan distinguir entre estados epistémicos que justifican acción y estados que exigen espera.

7. Núcleo técnico: cinco pilares de la metaciencia operatoria

La metaciencia operatoria se articula en torno a cinco pilares técnicos que, conjuntamente, permiten evaluar la robustez de un cuerpo de evidencia. Ninguno de estos pilares es suficiente por sí solo; es su integración sistemática lo que permite aproximarse a juicios de cierre operatorio.

7.1. Reproducibilidad y replicación

Qué es: La reproducibilidad refiere a la capacidad de obtener resultados idénticos aplicando los mismos métodos a los mismos datos. La replicación, más exigente, implica obtener resultados consistentes aplicando los mismos métodos a datos nuevos, idealmente recolectados por investigadores independientes.

Qué problema resuelve: La reproducibilidad detecta errores computacionales, transcripción incorrecta de datos y fraude directo. La replicación evalúa la generalidad de los hallazgos más allá de las condiciones específicas del estudio original, distinguiendo entre efectos robustos y artefactos locales.

Cómo se implementa: La reproducibilidad requiere compartir datos y código analítico. La Open Science Collaboration (2015) estableció protocolos de replicación que incluían colaboración con autores originales, preregistro de protocolos y análisis predefinidos. Plataformas como el Open Science Framework facilitan la transparencia necesaria.

Cómo puede fallar o ser manipulado: La replicación puede fracasar por razones legítimas: variabilidad muestral, diferencias contextuales, heterogeneidad de efectos. Los defensores de hallazgos fallidos pueden invocar estas explicaciones sistemáticamente, convirtiendo la replicación en un ejercicio indecible. Además, el sesgo de confirmación puede infiltrarse incluso en replications aparentemente preregistradas mediante decisiones post-hoc sobre exclusiones o análisis secundarios.

7.2. Preregistro y Registered Reports

Qué es: El preregistro implica especificar públicamente hipótesis, métodos y plan de análisis antes de recolectar datos. Los *Registered Reports* llevan esto más lejos: la revisión por pares ocurre antes de la recolección, y la publicación se garantiza independientemente de los resultados.

Qué problema resuelve: El preregistro combate el *HARKing* (*Hypothesizing After Results are Known*), la práctica de presentar predicciones post-hoc como si fueran a priori. También reduce la flexibilidad analítica documentada por Simmons, Nelson y Simonsohn (2011), al fijar las decisiones metodológicas antes de conocer su impacto en los resultados.

Cómo se implementa: Plataformas como AsPredicted y OSF Registries permiten preregistros con sellado temporal. Las revistas que aceptan *Registered Reports* evalúan la importancia de la pregunta y la solidez del diseño, eliminando el sesgo de publicación basado en resultados.

Cómo puede fallar o ser manipulado: Los preregistros pueden ser vagos, permitiendo flexibilidad interpretativa posterior. Los investigadores pueden registrar múltiples estudios y publicar selectivamente los exitosos. Además, el preregistro no garantiza calidad metodológica: un estudio puede estar perfectamente preregistrado y ser conceptualmente trivial o estadísticamente infrapotente.

7.3. Estadística con potencia, sesgos y flexibilidad analítica

Qué es: Este pilar abarca el uso riguroso de herramientas estadísticas, incluyendo análisis de potencia a priori, corrección por comparaciones múltiples, evaluación de sesgos de publicación mediante técnicas como el *funnel plot* y la *p-curve*, y transparencia sobre el espacio de decisiones analíticas.

Qué problema resuelve: Button et al. (2013) demostraron que la neurociencia típica operaba con potencias del 8-31%, garantizando que la mayoría de los estudios fracasarían en detectar efectos reales y que los positivos serían predominantemente inflados o falsos. El análisis de potencia previene este desperdicio. Las técnicas de detección de sesgo permiten evaluar si un cuerpo de literatura refleja efectos genuinos o artefactos selectivos.

Cómo se implementa: El análisis de potencia requiere especificar el tamaño de efecto mínimo de interés —decisión sustantiva, no meramente estadística—. Los metaanálisis deben incluir evaluación de heterogeneidad, análisis de moderadores y pruebas de sesgo. La especificación *multiverse* explora cómo varían los resultados bajo diferentes decisiones analíticas plausibles.

Cómo puede fallar o ser manipulado: Los análisis de potencia pueden inflarse asumiendo tamaños de efecto irrealistamente grandes. Los metaanálisis pueden incluir selectivamente estudios que apoyen conclusiones deseadas. Las técnicas de detección de sesgo tienen sus propios supuestos violables. Smaldino y McElreath (2016) mostraron que los incentivos

académicos pueden producir «selección natural de mala ciencia» incluso cuando los investigadores individuales actúan racionalmente.

7.4. Trazabilidad: datos, código y materiales

Qué es: La trazabilidad implica documentación exhaustiva que permita seguir la cadena desde los datos brutos hasta las conclusiones publicadas, incluyendo código analítico, scripts de procesamiento, materiales experimentales y decisiones de exclusión.

Qué problema resuelve: Sin trazabilidad, los errores permanecen ocultos y la verificación independiente es imposible. Muchos fracasos de replicación han revelado que los estudios originales contenían errores de programación, exclusiones no documentadas o discrepancias entre datos y análisis reportados.

Cómo se implementa: Repositorios como GitHub para código, OSF para materiales, y repositorios de datos institucionales o disciplinarios. Los *notebooks* computacionales (Jupyter, R Markdown) integran código y narrativa. Las licencias abiertas facilitan reutilización. Los identificadores persistentes (DOI) aseguran citabilidad.

Cómo puede fallar o ser manipulado: La documentación puede ser incompleta o intencionalmente oscurecida. Los datos pueden compartirse en formatos que dificultan reanálisis. El código puede ser correcto pero ilegible. Además, compartir datos plantea problemas de privacidad en investigación con participantes humanos, creando tensiones legítimas entre transparencia y protección.

7.5. Gobernanza científica e incentivos institucionales

Qué es: Este pilar reconoce que las prácticas científicas están moldeadas por estructuras de incentivos —promoción académica, financiación, prestigio— y que la mejora sostenible requiere reforma institucional, no solo buena voluntad individual.

Qué problema resuelve: Smaldino y McElreath (2016) demostraron formalmente que cuando los incentivos premian publicaciones positivas sobre replications y cantidad sobre calidad, la «selección natural» favorece prácticas que inflan hallazgos espurios. Ninguna reforma metodológica individual puede contrarrestar incentivos sistémicamente perversos.

Cómo se implementa: Las reformas incluyen: valorar replications en evaluaciones de carrera; adoptar métricas alternativas que premien rigor sobre impacto; implementar *Registered*

Reports como formato estándar; requerir transparencia como condición de financiación; crear incentivos para detección y corrección de errores.

Cómo puede fallar o ser manipulado: Las reformas institucionales pueden implementarse superficialmente (*box-ticking*) sin cambio real de prácticas. Las métricas alternativas pueden gaming-earse igual que las tradicionales. Las estructuras de poder existentes pueden resistir cambios que amenacen sus posiciones. Además, la fragmentación disciplinaria dificulta reformas coordinadas.

www.ateliervelazquez.com

8. Exponentes y corrientes principales

La metaciencia contemporánea se ha desarrollado a través de contribuciones específicas que, tomadas conjuntamente, configuran el campo. A continuación se presentan los aportes más significativos, no como biografías sino como contribuciones conceptuales y metodológicas.

John P. A. Ioannidis articuló el marco teórico que fundamenta la metaciencia crítica. Su artículo de 2005 demostró mediante simulación y argumentación que, bajo condiciones plausibles (baja potencia, alta flexibilidad analítica, sesgo de publicación), la mayoría de los hallazgos publicados son probablemente falsos. Este resultado no era meramente pesimista sino diagnóstico: identificaba los parámetros que determinan la fiabilidad de una literatura. Ioannidis también introdujo métricas como el *Positive Predictive Value* de hallazgos científicos, traduciendo intuiciones informales sobre calidad en indicadores cuantificables.

Brian Nosek y la Open Science Collaboration transformaron las preocupaciones teóricas de Ioannidis en investigación empírica sistemática. El proyecto de replicación de 2015 no solo documentó tasas de replicación; estableció protocolos que se convirtieron en estándares del campo. Nosek además fundó el Center for Open Science y desarrolló infraestructura (OSF) que facilita prácticas abiertas. Su contribución es tanto metodológica como institucional: creó las herramientas y organizaciones necesarias para sostener la metaciencia como empresa colectiva.

Marcus Munafò y colaboradores sintetizaron las diversas líneas de crítica en el *Manifesto for Reproducible Science* (2017). Este documento organizaba las amenazas a la ciencia fiable y las soluciones propuestas en un marco coherente. Munafò también contribuyó al desarrollo de técnicas como la *p-curve* para detectar sesgo de publicación y al análisis de cómo las prácticas de investigación varían entre disciplinas.

Colin Camerer y colaboradores extendieron el paradigma de replicación a la economía experimental. Su estudio de 2018 replicó 21 estudios publicados en *Nature* y *Science*, encontrando que aproximadamente el 60% replicaban pero con efectos significativamente menores. Este trabajo demostró que los problemas de reproducibilidad no eran específicos de la psicología sino sistémicos en las ciencias sociales cuantitativas.

Joseph Simmons, Leif Nelson y Uri Simonsohn aportaron la demostración más pedagógica de cómo la flexibilidad analítica infla falsos positivos. Su artículo de 2011 mostró

experimentalmente que decisiones aparentemente inocuas sobre exclusión de datos, covariables y definiciones de variables podían producir cualquier resultado deseado. Además, propusieron soluciones concretas —preregistro, transparencia, 21-word solution— que se incorporaron rápidamente a las normas del campo.

Katherine Button y colaboradores publicaron en 2013 el análisis definitivo sobre potencia estadística en neurociencia. Demostraron que la potencia mediana era del 21%, muy por debajo del 80% convencionalmente aceptable. Este resultado explicaba por qué los efectos reportados estaban sistemáticamente inflados y por qué la literatura contenía inevitablemente altas tasas de falsos positivos.

C. Glenn Begley y Lee Ellis proporcionaron la perspectiva de la industria farmacéutica. Su reporte de 2012 sobre el fracaso en reproducir estudios preclínicos «landmark» en oncología fue devastador precisamente porque provenía de actores con interés comercial en replicar hallazgos prometedores. Su trabajo subrayó que los problemas de reproducibilidad no eran meramente académicos sino que tenían consecuencias directas para el desarrollo de tratamientos.

Paul Smaldino y Richard McElreath introdujeron la perspectiva evolutiva en la metaciencia. Su modelo de 2016 demostró que los incentivos académicos actuales producen «selección natural de mala ciencia»: los grupos que adoptan métodos que inflan hallazgos obtienen más publicaciones, financiación y estudiantes, propagando estas prácticas incluso sin intención de fraude. Esta contribución desplazó el foco de la responsabilidad individual hacia las estructuras sistémicas.

9. El Materialismo Filosófico como marco de cierre

9.1. Pertinencia del Materialismo Filosófico

Los cinco pilares técnicos expuestos en la sección anterior proporcionan herramientas para evaluar aspectos específicos de la práctica científica, pero carecen de un criterio integrador que permita emitir juicios de cierre operatorio. ¿Cuánta reproducibilidad es suficiente? ¿Qué nivel de potencia justifica actuar sobre los resultados? ¿Cómo ponderar transparencia metodológica contra otros valores? Estas preguntas no pueden responderse desde dentro de las herramientas técnicas; requieren un marco filosófico que articule qué significa para una ciencia alcanzar clausura operativa.

El Materialismo Filosófico desarrollado por Gustavo Bueno ofrece precisamente este marco. A diferencia de otras filosofías de la ciencia que operan en registros puramente normativos (falsacionismo popperiano) o puramente descriptivos (sociología del conocimiento), el Materialismo Filosófico proporciona una *gnoseología materialista* que conecta las operaciones técnicas concretas de las ciencias con criterios de cierre categorial. Como señala Bueno (1993), una ciencia no se define por su objeto de estudio ni por su método general sino por la producción de *identidades sintéticas* mediante operaciones sobre campos materiales específicos.

9.2. Symploké y la estructura M1–M2–M3

El Materialismo Filosófico postula una ontología trigenérica que distingue tres géneros de materialidad: **M1** (materia física, corpórea, espaciotemporalmente determinada), **M2** (materialidad de los procesos psicológicos, intencionales, operatorios) y **M3** (materialidad de los objetos abstractos, relaciones lógico-matemáticas, estructuras formales). Esta tricotomía no es una clasificación convencional sino una *symploké* —entrelazamiento dialéctico— donde cada género es irreducible a los otros pero interdependiente con ellos.

La noción de *symploké* (Bueno, 1972) es crucial para la metaciencia operatoria. Contra el monismo que reduce todo a un solo género (fiscalismo que reduce M2 y M3 a M1; idealismo que reduce M1 a M3), la *symploké* sostiene que «no todo está conectado con todo ni nada está aislado de todo». Esta tesis tiene consecuencias directas para entender la ciencia: no existe una *scientia una* que abarque todos los dominios, pero las ciencias particulares tampoco son compartimentos estancos. Cada ciencia configura su campo operando sobre materiales específicos pero estableciendo conexiones con otros campos.

Para la metaciencia operatoria, la estructura M1–M2–M3 permite distinguir entre los componentes físicos de la investigación (aparatos, muestras, datos: M1), las operaciones de los investigadores (diseño experimental, análisis, interpretación: M2) y las estructuras formales que articulan los resultados (teorías, modelos, relaciones matemáticas: M3). Una ciencia alcanza cierre categorial cuando sus operaciones (M2) sobre materiales (M1) producen identidades sintéticas estables (M3) que no dependen del sujeto operatorio particular.

9.3. Evitación del reduccionismo doble

El Materialismo Filosófico evita dos reduccionismos simétricos que amenazan a la metaciencia. El **reduccionismo sociologista** (asociado con el «programa fuerte» de la sociología del conocimiento científico) reduce la ciencia a sus determinaciones sociales: intereses de clase, dinámicas de poder, construcción social de hechos. Desde esta perspectiva, la verdad científica es una categoría vacía; solo existen consensos contingentes moldeados por factores extracientíficos. Para la metaciencia operatoria, esta posición es autodestructiva: si toda ciencia es construcción social, la propia metaciencia carece de privilegio epistémico para diagnosticar fallas.

El **reduccionismo formalista**, en el extremo opuesto, reduce la ciencia a estructuras lógico-matemáticas desconectadas de operaciones materiales concretas. El positivismo lógico intentó reconstruir la ciencia como sistema axiomático; el falsacionismo popperiano privilegió las relaciones lógicas entre teorías y enunciados observacionales. Para la metaciencia operatoria, este reduccionismo es igualmente problemático porque ignora que las ciencias son prácticas materiales realizadas por sujetos concretos en contextos institucionales específicos.

El Materialismo Filosófico supera ambos reduccionismos mediante la teoría del cierre categorial. La ciencia no es ni pura construcción social ni sistema formal puro; es producción de identidades sintéticas mediante operaciones. Los factores sociológicos (incentivos, sesgos, dinámicas de grupo) afectan la ciencia pero no la constituyen; las estructuras formales expresan los resultados pero no los generan. La metaciencia operatoria debe atender a los tres géneros de materialidad sin colapsar ninguno en los otros.

9.4. Cierre categorial como criterio de científicidad efectiva

El concepto de *cierre categorial* (Bueno, 1993; Casado, 2024) proporciona el criterio de científicidad que la metaciencia operatoria requiere. Una ciencia alcanza cierre cuando sus operaciones sobre un campo material producen términos que permanecen dentro del campo,

estableciendo *identidades sintéticas* que no dependen de las circunstancias particulares de su producción. La geometría euclidiana es paradigmática: las operaciones (construcciones con regla y compás) sobre términos (puntos, líneas, círculos) producen relaciones (teoremas) que son verdaderas independientemente del geómetra particular.

El cierre no es una propiedad binaria (presente/ausente) sino gradual. Bueno (1995) distingue entre cierres más o menos firmes según la estabilidad de las identidades producidas. Las ciencias «duras» exhiben cierres más robustos porque sus identidades son menos dependientes de condiciones contextuales; las ciencias «blandas» operan con cierres más frágiles, susceptibles a variabilidad operatoria. Crucialmente, esta distinción no es valorativa: no dice que unas ciencias sean «mejores» que otras, sino que sus modos de cierre difieren.

Para la metaciencia operatoria, el cierre categorial ofrece el criterio integrador que los pilares técnicos carecían. La reproducibilidad es indicador de cierre: si las identidades producidas no dependen del sujeto operatorio, diferentes investigadores deberían obtener los mismos resultados. La potencia estadística es condición de cierre: solo con potencia adecuada pueden detectarse los efectos que constituyen las identidades del campo. El preregistro protege contra pseudo-cierres: impide que identidades espurias (artefactos del *p-hacking*) contaminen el campo. La trazabilidad garantiza verificabilidad del cierre. La gobernanza institucional crea condiciones para que el cierre ocurra.

La pregunta operatoria fundamental —¿cuándo está justificado actuar sobre evidencia científica?— encuentra así respuesta: cuando el campo ha alcanzado cierre suficiente para que las identidades producidas sean fiables guías de acción. Determinar este umbral requiere evaluación conjunta de los cinco pilares, interpretados a la luz del criterio de cierre categorial.

10. Mini-caso aplicado: Economía conductual

10.1. Importancia práctica del campo

La economía conductual (*behavioral economics*) ha ejercido influencia extraordinaria en políticas públicas durante las últimas dos décadas. Desde las «unidades nudge» instituidas en gobiernos de múltiples países hasta las aplicaciones en diseño de planes de pensiones, opciones de seguros médicos y comunicación de riesgos, los hallazgos de esta disciplina se han traducido directamente en intervenciones que afectan a millones de personas. La premisa fundamental —que los agentes económicos exhiben sesgos sistemáticos predecibles que pueden aprovecharse para mejorar decisiones— ha justificado una industria de consultoría y diseño de políticas.

Esta influencia práctica hace de la economía conductual un caso paradigmático para la metaciencia operatoria. No se trata de debates académicos abstractos: las decisiones sobre ahorro para la jubilación, selección de tratamientos médicos o cumplimiento tributario dependen de la solidez de hallazgos conductuales específicos. Si los efectos documentados son artefactos metodológicos o inflaciones estadísticas, las políticas basadas en ellos podrían ser ineficaces o incluso contraproducentes.

10.2. Evidencia disponible

La economía conductual descansa sobre un catálogo de sesgos y heurísticas documentados experimentalmente: aversión a la pérdida, efectos de encuadre, descuento hiperbólico, sesgo del *status quo*, efecto dotación, entre muchos otros. Los trabajos fundacionales de Kahneman y Tversky establecieron el paradigma mediante experimentos de laboratorio con estudiantes universitarios, típicamente enfrentados a decisiones hipotéticas sobre loterías y apuestas.

Metaanálisis recientes han cuantificado los tamaños de efecto promedio para los sesgos principales. La aversión a la pérdida, por ejemplo, se estima tradicionalmente en un ratio de 2:1 (las pérdidas pesan aproximadamente el doble que las ganancias equivalentes), aunque estimaciones más recientes sugieren valores más modestos. Los efectos de encuadre exhiben heterogeneidad considerable según el dominio de aplicación. Los efectos de los «nudges» en estudios de campo muestran tamaños sustancialmente menores que los experimentos de laboratorio originales.

El estudio de replicación de Camerer et al. (2018) incluyó varios experimentos clásicos de economía conductual. Los resultados fueron mixtos: algunos efectos replicaron robustamente (aversión a la pérdida en ciertos paradigmas), mientras que otros exhibieron magnitudes reducidas o no replicaron en absoluto. Crucialmente, incluso los efectos replicados mostraron tamaños consistentemente menores que los estudios originales, sugiriendo inflación sistemática en la literatura publicada.

10.3. Fallas operatorias detectables

Aplicando los cinco pilares de la metaciencia operatoria al campo de la economía conductual, emergen varias fallas detectables. Respecto a **reproducibilidad y replicación**, la literatura presenta patrones preocupantes: muchos estudios clásicos no han sido replicados sistemáticamente; los que sí han sido replicados muestran tasas de éxito modestas; los tamaños de efecto disminuyen consistentemente en replications, indicando inflación en estudios originales.

Respecto al **preregistro**, la adopción en economía conductual ha sido más lenta que en psicología. Muchos de los hallazgos fundacionales del campo predatan la era del preregistro y nunca han sido sometidos a confirmación mediante protocolos preregistrados. Los estudios recientes con preregistro muestran tamaños de efecto consistentemente menores que la literatura histórica.

Respecto a **potencia estadística**, los estudios de laboratorio típicos operaron con muestras pequeñas ($N < 50$ por condición), garantizando potencia inadecuada para efectos modestos. Los análisis de potencia retrospectivos sugieren que gran parte de la literatura está infrapotenciada, lo que explicaría tanto la inflación de efectos como la variabilidad entre replications.

Respecto a **trazabilidad**, muchos estudios clásicos carecen de datos y materiales compartidos, imposibilitando reproducibilidad computacional. Las prácticas de transparencia han mejorado en publicaciones recientes, pero el corpus histórico permanece opaco.

Respecto a **gobernanza e incentivos**, el campo exhibe las mismas presiones que otras disciplinas académicas: premio a la novedad sobre la replicación, sesgo de publicación hacia resultados positivos, incentivos para hallazgos «sorprendentes» que atraigan cobertura mediática. El éxito comercial de la economía conductual ha añadido presiones adicionales:

consultores e implementadores de políticas tienen incentivos para sobreestimar la efectividad de las intervenciones.

10.4. Dictamen: cautela con apertura selectiva

Integrando la evaluación de los cinco pilares bajo el criterio de cierre categorial, el dictamen para la economía conductual es de **cautela con apertura selectiva**. No corresponde ni rechazo total (los efectos conductuales existen; la cuestión es su magnitud y generalidad) ni adopción incondicional (la evidencia está inflada y la heterogeneidad es sustancial).

Específicamente, la economía conductual exhibe un cierre parcial e inestable. Algunas identidades —como la existencia de efectos de encuadre y aversión a la pérdida— parecen suficientemente robustas para justificar consideración en diseño de políticas. Sin embargo, los tamaños de efecto específicos y las condiciones de aplicabilidad permanecen indeterminadas. Actuar como si el ratio 2:1 de aversión a la pérdida fuera universal sería imprudente; reconocer que las pérdidas tienden a pesar más que las ganancias es razonable.

Las intervenciones de política basadas en economía conductual deberían, por tanto, tratarse como experimentos más que como implementaciones de conocimiento establecido. Esto implica: diseñar evaluaciones rigurosas de efectividad en contextos específicos; aceptar que los efectos de laboratorio probablemente sobrestiman los efectos de campo; mantener apertura a abandonar intervenciones que no demuestren eficacia robusta; evitar generalizar de dominios donde los efectos son robustos a dominios donde la evidencia es más débil.

10.5. Condiciones para considerar cierre suficiente

Para que la economía conductual alcanzara cierre suficiente que justificara implementación más confiada de políticas, serían necesarias las siguientes condiciones:

Primera condición: Replicaciones preregistradas de los efectos principales en muestras grandes y representativas, con tamaños de efecto estabilizados que no disminuyan sistemáticamente en cada nueva replicación.

Segunda condición: Metaanálisis con corrección de sesgo de publicación que proporcionen estimaciones descontadas de los tamaños de efecto reales, junto con análisis de moderadores que especifiquen condiciones de aplicabilidad.

Tercera condición: Evaluaciones de campo de intervenciones conductuales con diseños experimentales rigurosos, seguimiento a largo plazo, y reporte transparente de efectos nulos o contraproducentes.

Cuarta condición: Desarrollo de teoría predictiva que vaya más allá del catálogo de sesgos hacia modelos integrados que especifiquen cuándo y por qué los efectos conductuales emergen, permitiendo predicción ex ante más que explicación post hoc.

Mientras estas condiciones no se satisfagan, la metaciencia operatoria recomienda tratar los hallazgos de economía conductual como hipótesis prometedoras más que como conocimiento consolidado, calibrando las intervenciones prácticas en consecuencia.

www.ateliervelazquez.com

11. Patologías: no-cierre y sobrecierre

La metaciencia operatoria debe reconocer que las fallas en la producción de conocimiento pueden ocurrir en dos direcciones opuestas. Así como existe el riesgo de actuar sobre evidencia insuficiente (no-cierre), existe también el riesgo de osificación prematura del conocimiento (sobrecierre). Ambas patologías tienen manifestaciones características y consecuencias dañinas.

11.1. No-cierre: fragmentación, contradicción y ruido

El **no-cierre** caracteriza campos donde las operaciones sobre el material no producen identidades sintéticas estables. Sus manifestaciones incluyen: resultados contradictorios entre estudios sin resolución metodológica; fragmentación en escuelas o paradigmas que operan con conceptos incompatibles; incapacidad de acumular conocimiento porque cada estudio parece empezar de cero; y predominio de «ruido» (variabilidad no sistemática) sobre «señal» (efectos replicables).

Gran parte de la crisis de reproducibilidad puede interpretarse como síntoma de no-cierre. Los campos donde las replicaciones fracasan sistemáticamente son campos donde las identidades postuladas no se han consolidado. Los efectos de priming social, por ejemplo, exhibieron no-cierre: las operaciones (manipulaciones experimentales) no producían consistentemente los efectos (cambios conductuales) que los estudios originales reportaban. Lo que parecía conocimiento establecido era, en realidad, fluctuación estadística amplificada por sesgos de publicación.

Las consecuencias prácticas del no-cierre son la parálisis decisional y el desperdicio de recursos. Si no podemos determinar qué es verdadero, no podemos actuar racionalmente. Las políticas basadas en campos en no-cierre son apuestas, no aplicaciones de conocimiento. Los recursos invertidos en investigación que no acumula se desperdician. La confianza pública en la ciencia se erosiona cuando hallazgos celebrados se desvanecen.

11.2. Sobrecierre: dogma, circularidad e inmunización

El **sobrecierre** es la patología simétrica: el campo se comporta como si hubiera alcanzado cierre definitivo cuando en realidad las identidades producidas son espurias o prematuras. Sus manifestaciones incluyen: tratamiento dogmático de hallazgos particulares como axiomas incuestionables; circularidad donde la teoría determina qué cuenta como

evidencia relevante; inmunización contra anomalías mediante hipótesis ad hoc; y resistencia institucional a revisiones que amenacen consensos establecidos.

El sobrecierre puede resultar de éxito pasado (campos que alcanzaron cierre legítimo en algunas áreas y extrapolan injustificadamente a otras), de presiones institucionales (donde desafiar consensos amenaza carreras y financiación), o de dinámicas sociológicas (cascadas de información donde los investigadores siguen a la mayoría sin evaluación independiente). El concepto de *ciencia normal* kuhniano describe algunas de estas dinámicas, aunque Kuhn las interpretaba más benévolamente que la metaciencia operatoria.

Las consecuencias del sobrecierre son igualmente dañinas que las del no-cierre. Los recursos se canalizan hacia paradigmas potencialmente erróneos. Las voces disidentes son marginadas institucionalmente. Los errores persisten porque cuestionarlos tiene costos prohibitivos. La ciencia pierde capacidad autocorrectiva precisamente porque el mecanismo corrector (cuestionamiento de consensos) está bloqueado.

11.3. Riesgos simétricos y equilibrio operatorio

La metaciencia operatoria debe navegar entre Escila del no-cierre y Caribdis del sobrecierre. Este equilibrio no puede especificarse algorítmicamente de manera general; depende de las características específicas de cada campo, los costos de error tipo I versus tipo II en aplicaciones prácticas, y la madurez de las técnicas disponibles.

Sin embargo, algunos principios orientadores emergen. El escepticismo hacia hallazgos nuevos debe calibrarse según la potencia metodológica del campo: en campos con buenas prácticas, los hallazgos merecen más confianza inicial. La deferencia hacia consensos establecidos debe calibrarse según la calidad de la evidencia que los sustenta: consensos basados en estudios infrapotenciados merecen menos deferencia que consensos basados en replicaciones robustas. Los desafíos a consensos deben evaluarse según sus méritos metodológicos, no según la conveniencia institucional de mantener el statu quo.

El criterio de cierre categorial ayuda aquí porque proporciona un estándar independiente: ni la mera acumulación de publicaciones ni el consenso institucional constituyen cierre; solo la producción de identidades sintéticas estables mediante operaciones verificables. Un campo puede tener miles de publicaciones y seguir en no-cierre; un consenso puede ser unánime y constituir sobrecierre si se basa en evidencia metodológicamente deficiente.

12. Conclusión

El presente ensayo ha desarrollado la noción de metaciencia operatoria como respuesta al déficit decisional de la metaciencia contemporánea. Mientras la metaciencia descriptiva ha documentado extensamente las fallas de la producción científica —tasas de replicación decepcionantes, sesgos de publicación, flexibilidad analítica, incentivos perversos—, ha permanecido incapaz de articular criterios claros de cuándo la evidencia es suficiente para actuar. Este déficit no es meramente técnico sino filosófico: requiere criterios de cierre que las herramientas empíricas de la metaciencia no pueden proporcionar por sí solas.

Los cinco pilares de la metaciencia operatoria —reproducibilidad, preregistro, potencia estadística, trazabilidad y gobernanza— proporcionan las herramientas técnicas necesarias para evaluar la robustez de cuerpos de evidencia específicos. Sin embargo, estos pilares requieren integración mediante un marco filosófico que determine qué constituye suficiencia. El Materialismo Filosófico de Gustavo Bueno, y específicamente su concepto de cierre categorial, ofrece este marco.

La teoría del cierre categorial permite distinguir entre estados epistémicos que justifican acción y estados que exigen cautela o pausa. Una ciencia alcanza cierre cuando sus operaciones sobre un campo material producen identidades sintéticas estables, independientes del sujeto operatorio particular. Los cinco pilares técnicos se interpretan entonces como indicadores y condiciones de cierre: la reproducibilidad señala independencia del sujeto; la potencia estadística garantiza detectabilidad de efectos genuinos; el preregistro protege contra pseudo-cierres; la trazabilidad permite verificación; la gobernanza crea condiciones institucionales para el cierre.

El análisis del caso de la economía conductual ilustra cómo la metaciencia operatoria puede guiar decisiones prácticas. El campo exhibe un cierre parcial e inestable: algunos efectos parecen robustos, pero los tamaños específicos y las condiciones de aplicabilidad permanecen indeterminados. El dictamen apropiado es cautela con apertura selectiva: las intervenciones de política deberían tratarse como experimentos más que como implementaciones de conocimiento establecido.

La identificación de las patologías simétricas de no-cierre y sobrecierre completa el marco conceptual. La metaciencia operatoria debe evitar tanto la parálisis del no-cierre (donde la incertidumbre impide toda acción) como el dogmatismo del sobrecierre (donde consensos

prematuras bloquean la autocorrección). El equilibrio entre estos extremos no admite especificación algorítmica universal; depende de las características de cada campo y los costos relativos de diferentes tipos de error.

Finalmente, este ensayo ha mantenido deliberadamente abierta la posibilidad de una *formalización computable de los criterios de cierre operatorio*. Si la metaciencia operatoria ha de ser verdaderamente operatoria, sus criterios deben eventualmente traducirse en procedimientos algorítmicos aplicables sistemáticamente. Tal formalización requeriría especificar umbrales cuantitativos para cada pilar, reglas de agregación entre pilares, y mecanismos de actualización ante nueva evidencia. El desarrollo de tales sistemas formales constituye una dirección natural para la investigación futura en este campo.

El Materialismo Filosófico proporciona el marco gnoseológico para esta empresa, pero su implementación concreta queda como tarea pendiente. Lo que este ensayo ha procurado establecer es la necesidad de tal empresa y las coordenadas conceptuales dentro de las cuales debe desarrollarse: ni reduccionismo sociologista que disuelve la verdad en construcción social, ni reduccionismo formalista que ignora las condiciones materiales de la producción científica, sino gnoseología materialista que articula operaciones, materiales y estructuras formales en criterios de cierre efectivamente operatorios.

13. Glosario

Cierre categorial: Concepto central de la gnoseología de Gustavo Bueno que designa el estado en que las operaciones de una ciencia sobre un campo material producen identidades sintéticas estables, independientes del sujeto operatorio particular. Constituye el criterio de cientificidad efectiva.

Flexibilidad analítica: Conjunto de decisiones metodológicas (exclusión de datos, selección de covariables, definición de variables, etc.) que el investigador puede tomar y que afectan los resultados. Su explotación sistemática infla tasas de falsos positivos.

Gobernanza científica: Estructuras institucionales (sistemas de promoción, criterios de financiación, normas editoriales) que moldean las prácticas investigativas. La metaciencia operatoria reconoce que las reformas metodológicas individuales son insuficientes sin reforma de la gobernanza.

HARKing: Acrónimo de «Hypothesizing After Results are Known». Práctica de presentar predicciones generadas después de observar los datos como si hubieran sido formuladas a priori, inflando la aparente confirmación de teorías.

Identidad sintética: Resultado de operaciones científicas que establece relaciones estables entre términos del campo, no reducible a observaciones individuales ni a deducciones analíticas. Los teoremas geométricos son paradigmáticos.

M1 (Materia primogenérica): En la ontología del Materialismo Filosófico, género de materialidad que comprende entidades físicas, corpóreas, espaciotemporalmente determinadas. Los datos, muestras y aparatos experimentales pertenecen a M1.

M2 (Materia segundogenérica): Género de materialidad que comprende procesos psicológicos, intencionales, operatorios. Las decisiones del investigador, los diseños experimentales y las interpretaciones pertenecen a M2.

M3 (Materia terciogenérica): Género de materialidad que comprende objetos abstractos, relaciones lógico-matemáticas, estructuras formales. Las teorías, modelos y relaciones matemáticas pertenecen a M3.

Metaciencia: Estudio sistemático de la ciencia mediante métodos científicos. Incluye análisis de reproducibilidad, detección de sesgos, evaluación de prácticas metodológicas y estudio de estructuras institucionales.

Metaciencia descriptiva: Rama de la metaciencia que documenta características de la práctica científica (tasas de replicación, prevalencia de sesgos, patrones de publicación) sin pronunciarse sobre criterios de acción.

Metaciencia operatoria: Disciplina de segundo orden que establece criterios de decisión bajo incertidumbre epistémica, determinando umbrales de suficiencia evidencial que justifican acción práctica, cautela o suspensión del juicio.

No-cierre: Patología epistémica caracterizada por fragmentación, contradicción y predominio de ruido. Los campos en no-cierre no acumulan conocimiento porque las operaciones no producen identidades estables.

P-hacking: Práctica de manipular análisis estadísticos hasta obtener resultados «significativos» ($p < 0.05$), explotando la flexibilidad analítica para inflar falsos positivos.

Potencia estadística: Probabilidad de detectar un efecto verdadero de magnitud especificada si existe. Estudios infrapotenciados tienen alta probabilidad de falsos negativos y, cuando producen positivos, estos tienden a estar inflados.

Preregistro: Especificación pública de hipótesis, métodos y plan de análisis antes de recolectar datos. Reduce flexibilidad analítica y HARKing al fijar decisiones metodológicas antes de conocer su impacto en resultados.

Registered Reports: Formato de publicación donde la revisión por pares ocurre antes de la recolección de datos y la publicación se garantiza independientemente de los resultados. Elimina sesgo de publicación.

Replicación: Obtención de resultados consistentes aplicando los mismos métodos a datos nuevos, idealmente por investigadores independientes. Evalúa la generalidad de hallazgos más allá de condiciones específicas.

Reproducibilidad: Obtención de resultados idénticos aplicando los mismos métodos a los mismos datos. Detecta errores computacionales y garantiza trazabilidad.

Sesgo de publicación: Tendencia sistemática a publicar estudios con resultados positivos o «significativos» y a no publicar estudios con resultados nulos, distorsionando el corpus de literatura disponible.

Sobrecierre: Patología epistémica caracterizada por tratamiento dogmático de hallazgos como axiomas incuestionables, circularidad y resistencia institucional a revisión. Los campos en sobrecierre han «cerrado» prematuramente.

Symploké: Principio ontológico del Materialismo Filosófico según el cual «no todo está conectado con todo ni nada está aislado de todo». Implica pluralismo ontológico irreducible e interdependencia parcial entre géneros de materialidad.

Trazabilidad: Documentación exhaustiva que permite seguir la cadena desde datos brutos hasta conclusiones publicadas, incluyendo código, materiales y decisiones de exclusión. Condición necesaria para verificación.

www.ateliervelazquez.com

14. Referencias

- Begley, C. G., & Ellis, L. M. (2012). Raise standards for preclinical cancer research. *Nature*, 483(7391), 531–533. <https://doi.org/10.1038/483531a>
- Bueno, G. (1972). *Ensayos materialistas*. Taurus.
- Bueno, G. (1993). *Teoría del cierre categorial* (Vol. 2). Pentalfa.
- Bueno, G. (1995). *¿Qué es la ciencia?* Pentalfa.
- Bueno, G. (1996). *El mito de la cultura*. Prensa Ibérica.
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., ... & Wu, H. (2018). Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015. *Nature Human Behaviour*, 2(9), 637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- Casado, C. M. M. (2024). *La teoría del cierre categorial: génesis, estructura y sentido*. Fundación Gustavo Bueno.
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Melo, A. H. (2019). *Introducción a la teoría del cierre categorial*. Dialnet.
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., ... & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 0021. <https://doi.org/10.1038/s41562-016-0021>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, 3(9), 160384. <https://doi.org/10.1098/rsos.160384>

Pérez Jara, J. (2017). *Gustavo Bueno: Una introducción a su pensamiento*. Fundación Gustavo Bueno.

www.ateliervelazquez.com