

FORMAS DE CONOCER

Edición científica

Veinticuatro estudios sobre evidencia, representación, medición y conocimiento

Juan Segurola

Version v1.0 · 22 September 2026

© 2026 Juan Segurola. All Rights Reserved.

Índice

KNO	Capítulo	Página
KNO-001	Una marca en una piedra	5
KNO-002	Contar antes de los números	13
KNO-003	De las marcas a los símbolos	20
KNO-004	De los símbolos a la escritura	28
KNO-005	Lenguaje oral y transmisión	38
KNO-006	Mapas y representación espacial	47
KNO-007	Calendarios y medida del tiempo	57
KNO-008	Instrumentos que amplían los sentidos	66
KNO-009	Archivos, bibliotecas y memoria colectiva	76
KNO-010	La clasificación como herramienta de conocimiento	84
KNO-011	Medir antes de teorizar	93
KNO-012	Experimentos y control de variables	101
KNO-013	Modelos, mapas y realidad	109
KNO-014	Causalidad y correlación	117
KNO-015	Incertidumbre y probabilidad	129
KNO-016	Cómo se construye la evidencia: integrando las herramientas	139
KNO-017	¿Son las matemáticas un lenguaje?	148
KNO-018	Imágenes, diagramas y visualización	157
KNO-019	Datos, tablas y registros	164
KNO-020	Revisión, consenso y desacuerdo	173
KNO-021	Ciencia abierta y reproducibilidad	183
KNO-022	La computación como memoria externa	190
KNO-023	Cómo los grupos saben más que los individuos	196

KNO	Capítulo	Página
KNO-024	Conocimiento perdido y sesgo de preservación	203

Prólogo

Cómo deja huellas el conocimiento

El conocimiento rara vez llega como acceso directo a aquello que queremos conocer. Trabajamos a partir de huellas: marcas sobre piedra, tradiciones orales, mapas, calendarios, lecturas instrumentales, archivos, tablas, experimentos, modelos y conjuntos de datos preservados. Cada huella transporta información, pero también posee una historia de producción, selección y pérdida. El problema central de este volumen no es solo qué sabemos, sino cómo una afirmación llega a estar justificada.

Los veinticuatro estudios KNO siguen ese problema desde las primeras marcas externas y las prácticas de conteo hasta la escritura, la representación espacial y la medición; después avanzan hacia la causalidad, la incertidumbre, los modelos, la visualización, los registros, la revisión, la reproducibilidad, la computación y la inteligencia colectiva. Los capítulos finales vuelven el método sobre el propio archivo: qué desaparece, qué sobrevive y cómo la preservación puede sesgar la imagen reconstruida a partir de lo que queda.

Los capítulos forman una ruta coherente, pero no una escalera de progreso inevitable. Comunidades distintas resolvieron problemas de información distintos con herramientas diferentes. Una técnica posterior no es automáticamente una respuesta mejor a una pregunta anterior, y un objeto conservado no es automáticamente el primer ejemplo de una práctica. A lo largo de la serie se separan observación e interpretación, identidad bibliográfica y apoyo a una afirmación, y modelo útil y realidad representada.

Esta edición científica conserva los argumentos técnicos, los límites evidenciales y las referencias del corpus KNO, presentándolos como un único volumen. La disciplina recurrente es fácil de formular y difícil de mantener: preguntar qué se observó, qué hubo que inferir, qué supuestos conectan ambas cosas, qué comprobaciones independientes existen y qué evidencia obligaría a cambiar la conclusión.

Leído en orden, KNO es un estudio de la maquinaria del conocimiento. Leído de forma selectiva, cada capítulo sigue siendo un examen independiente de una parte de esa maquinaria. En ambos usos, la incertidumbre se trata como información sobre el alcance de las pruebas, no como permiso para considerar igualmente plausibles todas las explicaciones.

KNO-001 · Una marca en una piedra

KNO · Formas de conocer

Juan Segurola

v0.2 · ES

Todos los derechos reservados

1. Una incisión todavía no es un símbolo

La pregunta

Un surco puede medirse, fotografiarse y fecharse. Su significado no se extrae de la superficie del mismo modo. ¿Qué cadena de pruebas permite pasar de una huella física a una marca intencional y, después —con mayor cautela—, a un posible signo? Una línea sobre un objeto duradero comienza como una modificación observable, no como un mensaje descifrado. La arqueología puede establecer material, geometría, superposición y fabricación antes de abordar la finalidad. Ese orden importa porque un corte intencional pudo ser práctico, decorativo, imitativo, lúdico o comunicativo; su conservación no permite elegir entre esas posibilidades.[1,3] En Trinil, Java, una concha de agua dulce del conjunto asociado a Homo erectus presenta un grabado geométrico realizado antes de la fosilización. El contexto tiene cientos de miles de años. El estudio establece con más fuerza la fabricación deliberada que su función: sus autores declaran expresamente que no puede determinarse el propósito o significado del grabado.[1] Ocre de la Edad de Piedra Media procedentes de Blombos y Klein Kliphuis llevan patrones incisos deliberadamente. Los estudios detallados distinguen el grabado pautado del raspado o molido usado para procesar pigmento. Sin embargo, ni siquiera un reticulado repetido contiene su propio diccionario. Calificarlo de simbólico es una interpretación apoyada por patrón y contexto, no una traducción de contenido semántico recuperado.[2,3,5] Una lasca de silcreta de los niveles de Blombos de hace aproximadamente 73.000 años conserva líneas rojas entrecruzadas. La microscopía y el análisis químico respaldan una aplicación deliberada con un crayón de ocre, y las líneas continuaban más allá del borde conservado. Esto identifica un acto de dibujo y parte de un diseño mayor; no recupera qué significaba para quien lo hizo.[4] La inferencia tiene, por tanto, capas. El análisis superficial puede mostrar que una marca es antigua; la fractura y la secuencia, cuándo se hizo; los experimentos, una herramienta y un gesto; la recurrencia puede respaldar un patrón deliberado. Intención, notación y significado compartido se alejan progresivamente de la observación directa y requieren cada vez más contexto.[1,3,4]

En qué punto estamos

Las marcas duraderas más antiguas amplían la historia de la producción deliberada de patrones, pero no vienen acompañadas de leyendas. La arqueología es más sólida cuando expone la secuencia de inferencias sin colapsarla: objeto, modificación, gesto deliberado, forma recurrente, posible convención. La marca es evidencia; el significado recuperado suele seguir siendo la pregunta abierta.

2. Marcas, muescas y números son afirmaciones distintas

La pregunta

Varias muescas paralelas parecen contables a una mirada moderna. Pero que puedan contarse no significa que fueran contadas, y contar no equivale a una notación numérica. ¿Qué rasgos observables distinguen decoración, agarre, propiedad, acción secuencial y registro de acontecimientos discretos? Las incisiones pueden proceder de cortar, raspar, afilar, agarrar, descarnar o marcar deliberadamente. Los perfiles microscópicos, la dirección de los trazos, la superposición y los cambios de herramienta pueden excluir alternativas. Identifican cómo se produjo una secuencia con más seguridad que por qué. Una fila regular evidencia primero una acción organizada, no necesariamente conteo.[3,9] Los experimentos con el hueso de cuervo de Zaskalnaya indican que quizá se añadieron dos muescas para regularizar visualmente una secuencia previa. El resultado respalda la atención al espaciado. No demuestra que cada muesca representara un animal, un día o una cantidad, ni aporta una regla recuperada para leer la secuencia.[8,9] Un hueso con muescas de Border Cave, de unos 44.000 años, contiene grupos realizados en momentos distintos y se ha interpretado como notacional. La producción secuencial hace plausible un registro externo. Sin embargo, se desconoce la categoría registrada y el objeto no revela por sí solo si las unidades eran numéricas, calendáricas, sociales o prácticas.[7,9] Las propuestas arqueológicas sobre la aparición de símbolos numéricos emplean etapas comprobables: agrupación perceptiva, marcas uno a uno, acumulación entre sesiones, formas convencionales y, finalmente, operaciones con signos. Es un modelo para ordenar pruebas, no una secuencia universal excavada. Comunidades distintas pudieron exteriorizar cantidades mediante materiales y trayectorias diferentes.[9] El hueso de Ishango concentra la tentación de interpretar en exceso. Es un objeto pequeño, con punta de cuarzo y grupos de muescas, hallado cerca del lago Eduardo y fechado habitualmente en unos veinte mil años. Se han propuesto lecturas aritméticas y calendáricas. Las muescas son reales; no existe una decodificación única demostrada.[10]

En qué punto estamos

Una fila de muescas puede ser intencional sin ser un registro de conteo, y ese registro puede seguir objetos sin constituir escritura. Los análisis más sólidos comprueban fabricación, secuencia y agrupación antes de proponer contenido. El número entra en el argumento solo cuando la evidencia respalda una correspondencia entre marcas y unidades, no porque hoy podamos contarlas.

3. Cómo comprueba la arqueología la intención

La microscopía convierte una marca en una secuencia de acontecimientos físicos. Se examinan sección del surco, bordes, estrías internas, partículas de pigmento, puntos de entrada y salida y cruces entre líneas. Estos rasgos distinguen una punta arrastrada de pigmento depositado y reconstruyen qué trazo precedió a otro sin asignar un mensaje a ninguno.[3,4] La réplica experimental aporta comparación, no recreación histórica. Se usan herramientas y gestos candidatos sobre materiales semejantes y se comparan las huellas resultantes con las arqueológicas. Una coincidencia estrecha respalda una hipótesis de fabricación y estima habilidad o esfuerzo. No establece que el motivo del experimentador fuese el del autor antiguo.[3,8] El contexto pregunta si el objeto marcado pertenece realmente al lugar que se le atribuye. Estratigrafía, remontajes, datación, alteración superficial y asociación con otros artefactos limitan edad y conducta. Un objeto llamativo sin procedencia segura aporta menos información histórica que otro modesto excavado dentro de una secuencia documentada.[1,2,5] La confianza aumenta cuando convergen observaciones independientes: la marca precede al enterramiento, sus trazos se secuenciaron deliberadamente, los experimentos

reproducen su morfología y aparecen patrones comparables en contextos relacionados. Las señales no son independientes si parten de un objeto mal identificado; por eso identidad y procedencia deben verificarse antes de interpretar.[3,4,9] Un buen análisis conserva también los casos inciertos. El estudio de los ocre de Blombos separó grabados claros, incisiones ambiguas y huellas explicables por procesamiento. En general, la evidencia conductual pleistocena se acumula de modo desigual, no como un paquete repentino. Clasificar con incertidumbre explícita informa más que forzar cada marca en una revolución cognitiva.[3,6]

En qué punto estamos

La arqueología no puede entrevistar a quien hizo la marca, pero sí disciplinar las preguntas dirigidas al objeto. La microscopía reconstruye gestos, los experimentos comparan causas y el contexto asegura tiempo y asociación. El significado es la inferencia más distante. La incertidumbre no es un fracaso: registra la distancia entre una huella conservada y un acto desaparecido.

4. La memoria externa cambia la tarea

La pregunta

Recordar internamente y consultar una marca estable pueden producir la misma respuesta, pero no son la misma operación cognitiva. ¿Qué cambia cuando parte de una tarea queda fuera del sistema nervioso, donde puede inspeccionarse, reorganizarse y compartirse? Las representaciones externas cambian una tarea al hacer perceptibles algunos estados. Un resultado intermedio escrito, una ficha ordenada o una posición marcada no necesitan mantenerse por completo en la memoria de trabajo. El entorno no resuelve solo el problema; redistribuye almacenamiento, búsqueda y comparación entre una persona y una superficie estructurada.[15,18] En el estudio de cabina de Hutchins, las velocidades recordadas se distribuían entre personas, procedimientos hablados, instrumentos y marcadores móviles. Ningún componente contenía por separado toda la memoria. El caso demuestra cómo un sistema coordinado conserva estados relevantes, pero solo ofrece una analogía —no evidencia directa— para prácticas antiguas de marcar y comprobar.[16] Algunas acciones son epistémicas: se realizan para revelar, simplificar o estabilizar información, no para avanzar directamente hacia una meta física. Girar una pieza para reconocerla u ordenar fichas para comparar grupos reduce la búsqueda mental. Una marca puede cumplir ese papel aunque sea temporal y solo la entienda una persona.[17,18] La tesis de la mente extendida va más lejos: un recurso externo integrado de forma fiable podría contar como parte del proceso cognitivo. Es una posición filosófica sobre dónde termina la cognición, no una medición arqueológica. KNO-001 solo necesita el punto empírico más estrecho: los registros externos pueden cambiar cómo se recuerda y razona.[19] Para la prehistoria, la inferencia segura es condicional. Si las marcas se revisaban, correspondían una a una con acontecimientos o se acumulaban entre sesiones, pudieron reducir la carga del recuerdo sin ayuda. La secuencia microscópica respalda producción reiterada, pero la función cognitiva sigue siendo un modelo hasta vincular el objeto con una tarea recurrente.[9,15,18]

En qué punto estamos

Una marca externa no piensa, pero puede transformar lo que una persona debe mantener activo. Los estados estables pueden revisarse, compararse y corregirse; una disposición hace visibles relaciones. Los estudios cognitivos modernos establecen esos mecanismos. Aplicarlos a la prehistoria exige un argumento arqueológico independiente de que las marcas se usaron realmente en una tarea recurrente.

5. De la huella privada a la convención compartida

Una persona puede usar una marca sin que nadie más conozca su regla. Esa huella privada ya prolonga la persistencia durante un intervalo, pero deja de ser utilizable cuando se olvida la regla. La notación pública exige más: una comunidad debe estabilizar cómo se producen y distinguen las formas y con qué cosas o acciones se relacionan.[15,18] La recurrencia puede respaldar una convención, pero el parecido no basta. Retículas, líneas paralelas y cruces sencillas pueden aparecer de forma independiente porque manos y superficies limitan el diseño. Una prueba más fuerte combinaría forma repetida con fabricación, ubicación, cronología y contexto compartidos, admitiendo que signos semejantes pueden tener significados distintos.[3,4,5] La evolución cultural acumulativa implica modificaciones que se conservan y se apoyan en anteriores, aunque se distinguen varias formas y mecanismos. Los registros duraderos pueden ayudar a acumular al preservar resultados para aprendices posteriores. No son necesarios en todos los casos ni suficientes: un error persistente puede copiarse con tanta eficacia como una mejora.[25] Se ha propuesto el almacenamiento simbólico externo como gran transición de la historia cognitivo-cultural porque los registros sobreviven a la memoria individual y facilitan enseñanza, recuperación y revisión. El mecanismo general es plausible. Pero toda secuencia fija comprime historias diversas y debe contrastarse con cronologías arqueológicas, no asumirse como escalera universal.[20,25] Probablemente no hubo un instante en que una sola incisión convirtiera una mente aislada en cultura acumulativa. Patronaje, enseñanza, ayudas materiales, convención e instituciones dejan huellas distintas y no tienen por qué surgir juntas. Una historia mejor sigue cada capacidad y práctica por separado y pregunta cuándo ciertas combinaciones estabilizaron la transmisión.[6,20,25]

En qué punto estamos

La persistencia adquiere poder social cuando una regla sobrevive junto con la huella. Otra persona puede entonces recuperar, comparar o corregir lo registrado. La convención no es decoración añadida a la memoria: es la coordinación que hace transportable un estado externo entre mentes. La arqueología debe demostrar esa coordinación, no deducirla automáticamente de la geometría.

6. Contar objetos, llevar cuentas y escribir

La pregunta

Una ficha puede representar una oveja, una muesca una entrega y un signo escrito una palabra hablada. Estos sistemas pueden interactuar históricamente, pero resuelven problemas representacionales distintos. ¿Qué evidencia permite pasar de cantidad a cuenta y de cuenta a lenguaje? Humanos y otros animales discriminan numerosidades aproximadas, mientras los conjuntos pequeños exactos y los números simbólicos aprendidos implican sistemas parcialmente distintos. Los estudios neuronales identifican poblaciones sensibles a cantidad, pero sentir más o menos no es una inscripción. Los sistemas numéricos exactos e indefinidamente extensibles requieren prácticas aprendidas que vinculen formas, palabras o gestos con valores ordenados.[11,14] Experimentos con participantes pirahã hallaron precisión en algunas correspondencias exactas, pero errores crecientes cuando había que retener o transformar cantidades sin palabras numéricas exactas. El resultado respalda el lenguaje numérico como tecnología cognitiva para representar con estabilidad y exactitud. No demuestra ausencia de pensamiento numérico ni de inteligencia general.[12] Niños indígenas australianos que hablaban lenguas con vocabularios numéricos restringidos rindieron bien en tareas numéricas no verbales. Leído junto con los resultados pirahã, esto impide afirmar que las palabras numéricas creen toda capacidad. Lenguaje y signos materiales estabilizan operaciones exactas sin ser

la fuente de toda discriminación cuantitativa.[13] En el antiguo Oriente Próximo, fichas de arcilla, envoltorios y tablillas numéricas forman una reconstrucción influyente del registro administrativo anterior y paralelo a la escritura temprana. Los documentos protocuneiformes de finales del cuarto milenio a. C. son sobre todo numéricos y administrativos. La continuidad entre ficha y signo es potente para ese caso regional, pero sus detalles y universalidad siguen debatidos.[21,22] La escritura no es idéntica a todo sistema duradero de signos. Representa convencionalmente lenguaje en un ámbito suficientemente amplio, mientras muescas, marcas de propiedad y notaciones numéricas pueden codificar dominios restringidos. Las historias comparadas identifican varias invenciones independientes de escritura. No es necesario forzarlas por una única cadena iniciada en muescas.[21,23]

En qué punto estamos

La cantidad puede percibirse sin símbolos; palabras o marcas estabilizan valores exactos; las cuentas organizan dominios restringidos; y la escritura representa lenguaje. Estas capacidades se solapan, pero no son intercambiables. Sus historias deben reconstruirse localmente, a partir de transiciones documentadas, y no ordenarse en una escalera universal de muesca, número y texto.

7. Lo que sobrevive no es lo que existió

Hueso y concha entran en el registro arqueológico por vías selectivas. Grietas de meteorización, abrasión, pisoteo, alteración química, transporte y fractura eliminan objetos y modifican superficies. El análisis tafonómico identifica algunos procesos y muestra por qué la muestra conservada no es un censo neutral de lo que se fabricó o marcó.[24] Los registros orgánicos son especialmente vulnerables. Madera, corteza, fibra, piel y materiales vegetales blandos pueden portar marcas o disposiciones, pero suelen descomponerse salvo en condiciones excepcionales. Los artefactos orgánicos de Border Cave informan precisamente porque esa conservación es rara. Piedra y hueso pueden dominar la evidencia sin haber dominado la práctica antigua.[7,24] Un objeto duradero grabado tiene más probabilidades de conservarse, reconocerse y publicarse que una ayuda perecedera o sin marcas. La frecuencia en museos combina conducta antigua, conservación, excavación, identificación y atención investigadora. Las aparentes explosiones de marcado pueden mezclar cambio real y mayor visibilidad.[1,2,3,24] La ausencia de marcas en un nivel es más informativa cuando excavación, conservación y tamaño muestral permiten detectarlas. De otro modo puede indicar inexistencia de la práctica, otro material, otro contexto de uso o simple pérdida. El silencio arqueológico solo es evidencia después de modelar las condiciones capaces de producirlo.[6,24] Arqueología, ciencia cognitiva y neurociencia responden preguntas distintas. Las huellas limitan lo ocurrido a los objetos; los experimentos revelan capacidades y mecanismos de tarea; los datos neurales describen sistemas vivos. Pueden informarse entre sí, pero una capacidad moderna no fecha su origen y una marca antigua no revela toda la cognición que la produjo.[9,11,15,24]

En qué punto estamos

El archivo está filtrado antes de que comience la excavación. Materiales duraderos, contextos estables y patrones reconocibles están sobrerrepresentados; ayudas perecederas y experimentos fallidos desaparecen. Toda historia de la memoria externa debe explicar las marcas conservadas y los procesos que las seleccionaron. Un vacío del registro no equivale automáticamente a un vacío de capacidad.

8. Una marca cambia lo que puede persistir

Cuando un estado se estabiliza materialmente, puede persistir mientras la atención se desplaza. Quien lo produjo puede volver a él; otra persona quizá lo inspeccione; una acción posterior puede partir de un resultado anterior. Esta es la transformación mínima de la memoria externa, sea la marca artística, numérica o lingüística.[15,18] Un registro persistente permite además operaciones imposibles sobre un pensamiento fugaz: comparación simultánea, reordenación, anotación y corrección. Esas posibilidades hacen visibles discrepancias y distribuyen trabajo complejo en el tiempo. No garantizan verdad. Una entrada falsa duradera puede coordinar acciones posteriores con tanta eficacia como una correcta. [15,17,18] Cuando los registros adquieren reglas compartidas, prácticas de copia e instituciones, conectan a personas que nunca coinciden. Las cuentas coordinan existencias y obligaciones; los textos conservan procedimientos y argumentos; los archivos permiten comparaciones posteriores. Son expansiones históricamente documentadas de la persistencia, pero no deben proyectarse sobre toda línea paleolítica.[20,21,23,25] Ningún objeto conocido merece llamarse la primera memoria externa única. Las huellas anteriores pueden haberse perdido y los candidatos conservados expresan conductas distintas. La pregunta histórica es plural: ¿cuándo empezaron comunidades concretas a revisar marcas, vincularlas con unidades, compartir sus reglas y construir registros posteriores sobre ellas?[1,6,9,20] El cambio profundo no es que la piedra empezara a contener pensamiento, sino que una acción humana podía dejar un estado estructurado disponible después de terminar. De esa capacidad modesta surgieron, bajo convenciones e instituciones concretas, múltiples formas de conteo, diagrama, contabilidad y escritura: tecnologías relacionadas con historias distintas, no una escalera inevitable.[9,15,23,25]

En qué punto estamos

Una marca importa porque puede seguir disponible después del gesto y quizá después de su autor. Esa persistencia abre oportunidades de comprobación, intercambio y acumulación, pero cada una exige evidencia y práctica social adicionales. La historia científicamente honesta no salta de la incisión a la civilización: reconstruye ramificaciones de soportes externos para pensar.

Lo que sabemos y cómo lo sabemos

Microscopía, química, experimentos y contexto establecen que muchas marcas antiguas fueron producidas deliberadamente y a menudo reconstruyen herramientas, gestos y secuencia. Los experimentos cognitivos muestran que los estados externos redistribuyen memoria y comparación. Los sistemas administrativos y de escritura documentan usos convencionales posteriores. Estas conclusiones proceden de pruebas diferentes y deben mantenerse separadas.

Lo que todavía no sabemos

En la mayoría de los objetos tempranos ignoramos qué significaba una marca, si alguien volvía a consultarla, si su regla era compartida o si sus unidades eran numéricas. Tampoco podemos medir el registro desaparecido sobre madera, fibra, piel o superficies temporales. El ejemplo conservado más antiguo no es, por tanto, necesariamente el origen de la práctica.

Qué podría hacernos cambiar de opinión

Objetos marcados repetidos en contextos seguros, con secuencias de producción demostrables y asociaciones estables con tareas conocidas por otras vías, reforzarían las afirmaciones de notación o convención. Nuevos remontajes, dataciones o análisis de huellas también podrían debilitarlas. Una conservación excepcional de registros orgánicos puede revelar rutas hacia la memoria externa que hoy ocultan las colecciones duraderas.

Si solo recuerdas una cosa

Una marca deliberada conserva un estado más allá de un cerebro, pero símbolo, muesca, número y escritura exigen pruebas adicionales.

Artículos relacionados

Ninguno asignado en el Master Index v2.50.

Referencias

- [1] Josephine C. A. Joordens; Francesco d'Errico; Frank P. Wesselingh; Stephen Munro; John de Vos; Jakob Wallinga; Christina Ankjærgaard; Tony Reimann; Jan R. Wijbrans; Klaudia F. Kuiper; Herman J. Mûcher; Hélène Coqueugniot; Vincent Prié; Ineke Joosten; Bertil van Os; Anne S. Schulp; Michel Paniel; Victoria van der Haas; Wim Lustenhouwer; John J. G. Reijmer; Wil Roebroeks. (2015). Homo erectus at Trinil on Java used shells for tool production and engraving. *Nature* 518, 228–231. DOI: 10.1038/nature13962. <https://doi.org/10.1038/nature13962> [2] Christopher S. Henshilwood; Francesco d'Errico; Royden Yates; Zenobia Jacobs; Chantal Tribolo; Geoff A. T. Duller; Norbert Mercier; Judith C. Sealy; Hélène Valladas; Ian Watts; Ann G. Wintle. (2002). Emergence of modern human behavior: Middle Stone Age engravings from South Africa. *Science* 295, 1278–1280. DOI: 10.1126/science.1067575. <https://doi.org/10.1126/science.1067575> [3] Christopher S. Henshilwood; Francesco d'Errico; Ian Watts. (2009). Engraved ochres from the Middle Stone Age levels at Blombos Cave, South Africa. *Journal of Human Evolution* 57, 27–47. DOI: 10.1016/j.jhevol.2009.01.005. <https://doi.org/10.1016/j.jhevol.2009.01.005> [4] Christopher S. Henshilwood; Francesco d'Errico; Karen L. van Niekerk; Laure Dayet; Alain Queffelec; Luca Pollarolo. (2018). An abstract drawing from the 73,000-year-old levels at Blombos Cave, South Africa. *Nature* 562, 115–118. DOI: 10.1038/s41586-018-0514-3. <https://doi.org/10.1038/s41586-018-0514-3> [5] Alex Mackay; Aara Welz. (2008). Engraved ochre from a Middle Stone Age context at Klein Kliphuis in the Western Cape of South Africa. *Journal of Archaeological Science* 35, 1521–1532. DOI: 10.1016/j.jas.2007.10.015. <https://doi.org/10.1016/j.jas.2007.10.015> [6] Sally McBrearty; Alison S. Brooks. (2000). The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution* 39, 453–563. DOI: 10.1006/jhev.2000.0435. <https://doi.org/10.1006/jhev.2000.0435> [7] Francesco d'Errico; Lucinda Backwell; Paola Villa; Ilaria Degano; Jeannette J. Lucejko; Marion K. Bamford; Thomas F. G. Higham; Maria Perla Colombini; Peter B. Beaumont. (2012). Early evidence of San material culture represented by organic artifacts from Border Cave, South Africa. *Proceedings of the National Academy of Sciences* 109, 13214–13219. DOI: 10.1073/pnas.1204213109. <https://doi.org/10.1073/pnas.1204213109> [8] Ana Majkić; Sarah Evans; Vadim Stepanchuk; Alexander Tsvelikh; Francesco d'Errico. (2017). A decorated raven bone from the Zaskalnaya VI (Kolosovskaya) Neanderthal site, Crimea. *PLOS ONE* 12, e0173435. DOI: 10.1371/journal.pone.0173435. <https://doi.org/10.1371/journal.pone.0173435> [9] Francesco d'Errico; Luc Doyon; Ivan Colagè; Alain Queffelec; Emma Le Vraux; Giacomo Giacobini; Bernard Vandermeersch; Bruno Maureille. (2018). From number sense to number symbols: an

archaeological perspective. *Philosophical Transactions of the Royal Society B* 373, 20160518. DOI: 10.1098/rstb.2016.0518. <https://doi.org/10.1098/rstb.2016.0518> [10] Royal Belgian Institute of Natural Sciences. (Updated record accessed 2026). The Ishango Bone. Official collection/exhibition record. <https://www.naturalsciences.be/en/museum/exhibitions-activities/exhibitions/250-years-of-natural-sciences/the-ishango-bone> [11] Lisa Feigenson; Stanislas Dehaene; Elizabeth Spelke. (2004). Core systems of number. *Trends in Cognitive Sciences* 8, 307–314. DOI: 10.1016/j.tics.2004.05.002. <https://doi.org/10.1016/j.tics.2004.05.002> [12] Michael C. Frank; Daniel L. Everett; Evelina Fedorenko; Edward Gibson. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108, 819–824. DOI: 10.1016/j.cognition.2008.04.007. <https://doi.org/10.1016/j.cognition.2008.04.007> [13] Brian Butterworth; Robert Reeve; Fiona Reynolds; Delyth Lloyd. (2008). Numerical thought with and without words: Evidence from indigenous Australian children. *Proceedings of the National Academy of Sciences* 105, 13179–13184. DOI: 10.1073/pnas.0806045105. <https://doi.org/10.1073/pnas.0806045105> [14] Andreas Nieder. (2016). The neuronal code for number. *Nature Reviews Neuroscience* 17, 366–382. DOI: 10.1038/nrn.2016.40. <https://doi.org/10.1038/nrn.2016.40> [15] Jiajie Zhang; Donald A. Norman. (1994). Representations in distributed cognitive tasks. *Cognitive Science* 18, 87–122. DOI: 10.1207/s15516709cog1801_3. https://doi.org/10.1207/s15516709cog1801_3 [16] Edwin Hutchins. (1995). How a cockpit remembers its speeds. *Cognitive Science* 19, 265–288. DOI: 10.1207/s15516709cog1903_1. https://doi.org/10.1207/s15516709cog1903_1 [17] David Kirsh; Paul Maglio. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science* 18, 513–549. DOI: 10.1207/s15516709cog1804_1. https://doi.org/10.1207/s15516709cog1804_1 [18] David Kirsh. (2010). Thinking with external representations. *AI & Society* 25, 441–454. DOI: 10.1007/s00146-010-0272-8. <https://doi.org/10.1007/s00146-010-0272-8> [19] Andy Clark; David Chalmers. (1998). The extended mind. *Analysis* 58, 7–19. DOI: 10.1093/analys/58.1.7. <https://doi.org/10.1093/analys/58.1.7> [20] Merlin Donald. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press. <https://www.hup.harvard.edu/books/9780674644847> [21] Hans J. Nissen; Peter Damerow; Robert K. Englund. (1993). *Archaic Bookkeeping: Early Writing and Techniques of Economic Administration in the Ancient Near East*. University of Chicago Press. <https://cdli.earth/publications/1657148> [22] Denise Schmandt-Besserat. (1992). *Before Writing, Volume I: From Counting to Cuneiform*. University of Texas Press. <https://utpress.utexas.edu/9781477325766/> [23] Christopher Woods; Emily Teeter; Geoff Emberling (editors). (2010). *Visible Language: Inventions of Writing in the Ancient Middle East and Beyond*. Oriental Institute Museum Publications 32. <https://isac.uchicago.edu/research/publications/oimp/oimp-32-visible-language-inventions-writing-ancient-middle-east-and> [24] Anna K. Behrensmeyer. (1978). Taphonomic and ecologic information from bone weathering. *Paleobiology* 4, 150–162. DOI: 10.1017/S0094837300005820. <https://doi.org/10.1017/S0094837300005820> [25] Alex Mesoudi; Alex Thornton. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B* 285, 20180712. DOI: 10.1098/rspb.2018.0712. <https://doi.org/10.1098/rspb.2018.0712>

KNO-002 · Contar antes de los números

KNO · Formas de conocer

Juan Seguro

v0.1 · ES

Todos los derechos reservados

1. Cantidad antes de un número exacto

La pregunta

Si una persona puede elegir la cesta más llena sin decir cuántos objetos contiene cada una, ¿qué conocimiento numérico está presente y cuál falta todavía? Humanos, bebés y numerosos animales comparan numerosidades aproximadas. El rendimiento depende de la proporción: distinguir ocho de dieciséis resulta más fácil que distinguir ocho de diez. Este sistema numérico aproximado representa magnitud con precisión limitada; no genera por sí solo un símbolo exacto para cada entero positivo. [2,3] Un segundo sistema sigue con gran precisión un número pequeño de objetos individuales, normalmente dentro de un límite estrecho. Seguir uno, dos o tres objetos no es la misma operación que estimar una multitud, y ninguno de los dos sistemas aporta por sí solo la regla sucesora que hace indefinidamente extensibles los números de conteo.[2] Experimentos con participantes mundurukú mostraron capacidad para suma y comparación aproximadas pese a un léxico exacto limitado. Estudios con participantes pirahã hallaron también buen rendimiento cuando era posible emparejar objetos directamente, pero menor precisión al conservar, transformar o reproducir cantidades exactas tras un intervalo. Las palabras numéricas exactas son herramientas que estabilizan ciertas tareas, no requisitos para toda percepción cuantitativa.[4-6] Estos experimentos no ordenan lenguas ni pueblos por sofisticación cognitiva. Aíslan exigencias de tarea. Si la colección sigue visible, una correspondencia directa puede resolverla; si desaparece, una palabra, gesto o disposición material convencional conserva un estado exacto. El contraste está entre recursos representacionales, no entre mentes con y sin cantidad.[5-7]

En qué punto estamos

Magnitud aproximada, seguimiento de conjuntos pequeños y número simbólico exacto son capacidades relacionadas pero distintas, y ninguno de los dos sistemas básicos aporta la regla del sucesor que hace el conteo indefinidamente extensible. Antes de preguntar cuándo una población «tenía números», conviene preguntar qué tarea cuantitativa podía resolver, con qué precisión y mediante qué recursos internos o externos. El contraste aislado es entre recursos representacionales, no entre mentes.

2. La correspondencia uno a uno todavía no es un numeral

La pregunta

¿Puede demostrarse que dos conjuntos son iguales sin identificar su número cardinal? La correspondencia uno a uno empareja cada miembro de una colección con uno de otra. Una persona puede asignar una piedra a cada animal que sale y retirar una cuando regresa. Si no sobra piedra ni queda animal sin emparejar, los conjuntos corresponden aunque nadie use una palabra equivalente a «treinta y siete».[1,13] El procedimiento conserva igualdad y detecta exceso o ausencia. No representa

necesariamente el total con independencia de los objetos ni explica que todas las disposiciones con el mismo cardinal compartan un número abstracto. La relación puede seguir siendo práctica y local: un marcador por animal, recipiente u obligación.[13,14] Una muesca añade persistencia. Si una marca se vincula deliberadamente a un acontecimiento, el registro puede consultarse después. Pero las marcas repetidas no son automáticamente muescas de conteo. La arqueología debe establecer producción intencional y una unidad verosímil antes de tratar la regularidad como notación cuantitativa.[1] Incluso una muesca auténtica no constituye necesariamente un sistema numeral. Sus marcas pueden ser acumulaciones intercambiables y no signos nombrados con valores convencionales. La muesca permite responder «¿cuántas unidades han ocurrido aquí?» volviendo a contar; un numeral como 37 codifica un valor dentro de una notación estructurada transferible entre contextos.[18]

En qué punto estamos

Emparejar, registrar unidades y nombrar un cardinal resuelven problemas distintos. La correspondencia asegura equivalencia sin palabra numérica; una muesca conserva unidades sin numeral portátil; un sistema numeral vuelve convencionales y reutilizables los valores. Las marcas repetidas no son automáticamente muescas de recuento, y hay que establecer producción deliberada y una correspondencia de unidad plausible antes de leer la regularidad como notación cuantitativa.

3. Dedos, cuerpos y agrupación

Los sistemas corporales convierten una anatomía estable en recorrido ordenado. Los dedos son cómodos por ser visibles, móviles y estar agrupados, pero también existen secuencias documentadas que recorren muñecas, codos, hombros, cabeza y el lado opuesto del cuerpo. La secuencia funciona como orden externo aunque sus «signos» se ejecuten en lugar de separarse del cuerpo.[11,12] Contar con los dedos no es uniforme. Las culturas difieren en mano inicial, dirección, orden de dedos y función cardinal, ordinal u operatoria de los gestos. Esas diferencias influyen en el procesamiento numérico, por lo que la mano es una herramienta representacional aprendida y no un código universal listo para ser leído.[11] Agrupar reduce la carga de largas secuencias unitarias. Las colecciones se organizan en pares, cinco, decenas, veintenas u otros paquetes; una unidad de orden superior representa varias inferiores. Las bases reflejan a menudo prácticas corporales o dispositivos materiales, pero ninguna base se deduce inevitablemente de la anatomía humana.[18,20] Una lista ordenada se convierte en tecnología de número exacto cuando quien aprende coordina varios principios. Cada elemento recibe un término; se conserva el orden; cambiar la disposición no altera el total; y la última palabra se entiende como cardinal del conjunto. Un niño puede recitar números antes de comprender este principio cardinal.[9,10] Cuando se comprende la relación sucesora, añadir un elemento conecta un conjunto con el número siguiente. Eso permite generalizar más allá de valores pequeños memorizados. El desarrollo infantil no reproduce directamente la historia de una sociedad, pero aclara qué relaciones conceptuales debe coordinar un sistema maduro.[9,10]

En qué punto estamos

El cuerpo aporta orden, memoria y agrupación sin imponer una notación universal, y ninguna base se sigue inevitablemente de la anatomía humana. El conteo exacto surge de coordinar una secuencia aprendida, objetos individuales y un total cardinal, no de disponer simplemente de diez dedos. La evidencia del desarrollo aclara qué relaciones debe coordinar un sistema maduro; no reconstruye la historia de ninguna sociedad.

4. Fichas, cuentas e instituciones

Los contadores materiales conservan cantidad aunque los objetos representados se desplacen. Una ficha se manipula, agrupa, transfiere o sella; su utilidad depende de una regla que la vincula con categoría, obligación o mercancía. La estabilidad material permite comprobar, pero la regla sigue siendo social.[13,16] En el sudoeste de Asia aparecen fichas de arcilla mucho antes que las primeras tablillas. Denise Schmandt-Besserat propuso un desarrollo prolongado desde fichas hasta envoltorios sellados, impresiones y escritura. La reconstrucción es influyente porque conecta formas materiales con necesidades administrativas, pero se debaten el grado de continuidad directa y su alcance explicativo.[16] Las tablillas protocuneiformes de finales del cuarto milenio a. C. ofrecen evidencia más firme de registro numérico institucional. Muchas consignan cantidades y mercancías mediante diversos sistemas metrológicos y numéricos. El valor de un signo podía depender de qué se contaba o medía; la misma forma gráfica no funcionaba necesariamente como número abstracto independiente del contexto.[15,17] No eran simples cuadernos de aritmética. Pertenecían a prácticas administrativas con personal formado, convenciones, archivos y transacciones recurrentes. Institución y notación evolucionaron juntas: un sistema duradero hacía comparables las cuentas y la contabilidad repetida ofrecía el entorno donde estandarizar signos.[15] Esta secuencia regional no debe convertirse en ley universal. Otras sociedades recurrieron a nudos, palos, secuencias corporales, fórmulas orales, numerales escritos y dispositivos mixtos. Tareas semejantes utilizan medios distintos y medios parecidos pueden servir tareas diferentes.[14,18]

En qué punto estamos

La evidencia del Próximo Oriente documenta una transición especialmente rica desde contabilidad material hasta administración numérica escrita. Demuestra qué podían hacer fichas y tablillas en esa historia, no que todo sistema numérico o toda escritura comenzase igual. Allí institución y notación se desarrollaron juntas, y el grado de continuidad directa entre fichas y escritura sigue debatido y no resuelto.

5. De cantidades a numerales y valor posicional

Un numeral es un signo o secuencia convencional que representa un número. Su eficacia depende de reglas: formas básicas, repetición, valor por posición, presencia de cero y combinación de unidades de distinto orden.[18] Los sistemas aditivos repiten o combinan signos de valor; los multiplicativo-aditivos expresan productos de unidades; los posicionales asignan valor también por lugar. Son categorías analíticas, no rangos evolutivos. Una notación puede ser eficaz para sus usos sin converger hacia el valor posicional decimal moderno.[18] El valor posicional comprime la representación porque pocas cifras expresan números arbitrariamente grandes cuando las posiciones representan potencias de una base. También plantea un problema: distinguir una posición vacía de la ausencia de notación. Los marcadores de posición y el cero numérico tienen historias relacionadas pero no idénticas.[18,20] Los numerales escritos hacen portátiles los valores. «37» describe animales, días o longitudes porque la secuencia se abstrae de la categoría. Pero los sistemas prácticos suelen conservar unidades, medidas y signos de mercancía, recordando que la abstracción se consigue por convención y no borrando el contexto automáticamente.[15,18] Ninguna notación es destino inevitable del conteo. Numeración romana, sistemas sexagesimales mesopotámicos, varillas chinas, notación vigesimal mesoamericana y valor posicional decimal indoarábigo organizan de forma distinta representación, cálculo y transmisión.[18]

En qué punto estamos

Los numerales son convenciones diseñadas. Sus formas codifican decisiones sobre agrupación, orden, posición y ausencia, y las categorías que las describen son analíticas, no rangos evolutivos. El valor posicional es poderoso, pero constituye una solución histórica entre muchas, no la definición del pensamiento numérico, y su abstracción se logra por convención y no borrando el contexto.

6. Medir no es contar

Contar asigna un número a unidades discretas bajo una regla que define qué cuenta como una. Medir compara una magnitud continua con una unidad elegida. Cinco recipientes se cuentan; cinco litros resultan de relacionar repetidamente un volumen con un patrón.[15,18] Procedimentalmente se entrelazan porque medir suele requerir conteo de iteraciones, subdivisiones o marcas instrumentales. Conceptualmente difieren. Un total contado depende de individuar objetos; una medida depende de unidad, calibración y precisión admisible.[15] Las unidades son logros sociales y técnicos. Un codo ligado al cuerpo se hace más reproducible mediante patrones materiales y procedimientos institucionales. La estandarización permite transportar resultados y sitúa autoridad en quien mantiene la referencia y resuelve discrepancias.[15] Una medida puede ser fraccionaria o aproximada aunque se cuenten exactamente las divisiones del instrumento. Comunicarla exige, además del numeral, clase de magnitud, unidad, método e incertidumbre adecuada a la comparación.[18] Confundir conteo y medición distorsiona la arqueología. Un grupo de signos puede registrar objetos discretos, capacidad de grano, tiempo de trabajo o conversión contable. Las marcas solas no deciden qué cantidad representaban.[15,17]

En qué punto estamos

Contar trata unidades discretas; medir compara con un patrón, y las unidades son logros sociales y técnicos antes que hechos dados. Las prácticas históricas los combinan, pero interpretar un registro exige identificar sistema de unidades y magnitud representada, no solo leer un número visible. Unos signos agrupados pueden corresponder a objetos, capacidad, tiempo de trabajo o una conversión contable, y la notación por sí sola no decide cuál.

7. Qué pueden y qué no pueden establecer las marcas arqueológicas

El análisis arqueológico establece a menudo material, fabricación, orden de trazos, espaciado y sesiones distintas. Esas observaciones apoyan acumulación deliberada, pero no recuperan categoría contada, etiquetas orales ni regla que vinculaba marca y unidad.[1] La regularidad admite decoración, agarre, práctica, propiedad, ritmo, juego, agenda o muesca. La interpretación numérica se fortalece si evidencia independiente vincula las marcas con un proceso contable recurrente, los grupos obedecen una regla demostrable o varios objetos relacionados comparten convención.[1] Quien observa hoy siempre puede contar marcas antiguas. Ese hecho pertenece a la notación del observador y no necesariamente a la práctica del fabricante. Interpretaciones de primos, calendarios o patrones aritméticos son vulnerables a selección retrospectiva porque pueden descubrirse muchas agrupaciones en un conjunto finito.[1,14] La analogía etnográfica genera hipótesis sobre el uso cuantitativo de cuerpos, objetos o marcas. No asigna una pieza antigua a una práctica viva sin conexión histórica. Soluciones semejantes surgen independientemente y objetos parecidos participan en convenciones distintas.[12-14] La conservación filtra también el registro. Cuerdas, palos, fibras,

gestos y secuencias habladas sobreviven peor que hueso o piedra. El objeto de apariencia contable más antiguo no es necesariamente la primera ayuda cuantitativa, y la ausencia de notación duradera no demuestra ausencia de práctica exacta.[13,14]

En qué punto estamos

La secuencia segura es huella, fabricación, recurrencia y contexto. Cantidad, unidad y significado entran solo cuando evidencia adicional respalda la correspondencia, y que un observador moderno pueda contar marcas antiguas pertenece a la notación de ese observador y no a la práctica de quien las hizo. La conservación filtra lo que sobrevive, así que el objeto duradero más antiguo no es necesariamente el primer auxiliar cuantitativo.

8. Rutas múltiples, no una escalera

Contar es una familia de prácticas coordinadas. La aproximación compara magnitudes; el seguimiento de conjuntos pequeños mantiene individuos; la correspondencia empareja colecciones; la muesca conserva acontecimientos; las palabras estabilizan valores ordenados; la agrupación crea unidades superiores; los numerales vuelven portátiles los valores; y la medición relaciona magnitudes con patrones.[2,9,13,18] Estas capacidades se refuerzan sin aparecer necesariamente en orden fijo. Una comunidad puede tener conteo oral exacto y poca notación duradera, medición sofisticada con numerales especializados o contabilidad material estrechamente unida a mercancías. El cambio depende de tareas, enseñanza, medios e instituciones.[14,15,18] Los dispositivos externos no se limitan a almacenar un concepto abstracto ya completo. Al hacer visibles agrupaciones y correspondencias, cambian qué comparaciones y operaciones resultan fáciles de aprender y ejecutar. Pero ningún objeto genera una regla numérica sin usuarios que conserven la correspondencia.[13,19] La diversidad de sistemas numéricos es evidencia y no ruido. Bases, secuencias corporales, clasificadores, valores de signo y medios de cálculo distintos muestran que la abstracción se construye mediante relaciones repetibles, no se copia de una notación universal.[11,12,18] La historia de origen más defendible no tiene un objeto inicial único. Las capacidades biológicas son antiguas; los sistemas simbólicos exactos son culturales; los registros materiales conservados son selectivos; y las transiciones documentadas difieren. La historia del número es cómo las comunidades hicieron relaciones cuantitativas lo bastante estables para comprobarlas, compartirlas y extenderlas.[1,2,14]

En qué punto estamos

No hubo una marcha necesaria desde incisión a numeral y aritmética moderna. Hubo problemas recurrentes de emparejamiento, memoria, agrupación, intercambio y medida, resueltos con combinaciones distintas de cuerpo, lenguaje, materiales e instituciones. La diversidad de sistemas numéricos es evidencia y no ruido, y el relato de origen más defendible no tiene un único objeto inicial.

Lo que sabemos y cómo lo sabemos

Los experimentos controlados distinguen magnitud aproximada, seguimiento de conjuntos pequeños y prácticas exactas aprendidas. Los estudios del desarrollo muestran que recitar precede a comprender cardinalidad. La comparación intercultural demuestra rendimiento exacto no verbal bajo ciertas condiciones y la importancia de palabras o dispositivos cuando hay que conservar y transformar cantidades. Registros arqueológicos e históricos documentan contabilidad material y sistemas numerales administrativos en contextos concretos.[2,4-10,15-18]

Lo que todavía no sabemos

Para la mayoría de las marcas tempranas ignoramos si representaban cantidad, qué unidades seguían, si la regla era compartida o si coexistían palabras numéricas. Tampoco reconstruimos la mayoría perdida de dispositivos percederos y gestos. La diversidad cognitiva actual no ofrece una cronología directa de cuándo comunidades prehistóricas concretas coordinaron palabras, objetos y signos exactos.

Qué podría hacernos cambiar de opinión

Objetos marcados repetidos en contextos seguros, asociados a procesos contables conocidos por vías independientes y gobernados por una misma regla de agrupación demostrable, fortalecerían interpretaciones numéricas tempranas. Residuos, huellas de uso o remontajes podrían debilitarlas. Registros intermedios descifrables aclararían transiciones entre fichas, impresiones y numerales escritos. Investigación cognitiva más amplia puede revisar cómo interactúan sistemas aproximados, exactos y corporales.

Si solo recuerdas una cosa

Es posible manejar cantidad sin numerales, pero contar con exactitud exige correspondencia aprendida entre unidades, orden y cardinal; una marca repetida no demuestra por sí sola ninguna de esas relaciones.

Artículos relacionados

KNO-001 · Una marca en una piedra KNO-003 · De las marcas a los símbolos KNO-004 · De los símbolos a la escritura KNO-007 · Calendarios y medición del tiempo KNO-011 · Medir antes de teorizar

Referencias

- [1] Francesco d'Errico; Luc Doyon; Ivan Colagè; Alain Queffelec; Emma Le Vraux; Giacomo Giacobini; Bernard Vandermeersch; Bruno Maureille. (2018). From number sense to number symbols: an archaeological perspective. *Philosophical Transactions of the Royal Society B* 373, 20160518. DOI: 10.1098/rstb.2016.0518. <https://doi.org/10.1098/rstb.2016.0518> [2] Lisa Feigenson; Stanislas Dehaene; Elizabeth Spelke. (2004). Core systems of number. *Trends in Cognitive Sciences* 8, 307-314. DOI: 10.1016/j.tics.2004.05.002. <https://doi.org/10.1016/j.tics.2004.05.002> [3] Andreas Nieder. (2016). The neuronal code for number. *Nature Reviews Neuroscience* 17, 366-382. DOI: 10.1038/nrn.2016.40. <https://doi.org/10.1038/nrn.2016.40> [4] Peter Gordon. (2004). Numerical cognition without words: Evidence from Amazonia. *Science* 306, 496-499. DOI: 10.1126/science.1094492. <https://doi.org/10.1126/science.1094492> [5] Pierre Pica; Cathy Lemer; Véronique Izard; Stanislas Dehaene. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499-503. DOI: 10.1126/science.1102085. <https://doi.org/10.1126/science.1102085> [6] Michael C. Frank; Daniel L. Everett; Evelina Fedorenko; Edward Gibson. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108, 819-824. DOI: 10.1016/j.cognition.2008.04.007. <https://doi.org/10.1016/j.cognition.2008.04.007> [7] Brian Butterworth; Robert Reeve; Fiona Reynolds; Delyth Lloyd. (2008). Numerical thought with and without words: Evidence from indigenous Australian children.

Proceedings of the National Academy of Sciences 105, 13179-13184. DOI: 10.1073/pnas.0806045105. <https://doi.org/10.1073/pnas.0806045105> [8] Elizabet Spaepen; Marie Coppola; Elizabeth S. Spelke; Susan E. Carey; Susan Goldin-Meadow. (2011). Number without a language model. Proceedings of the National Academy of Sciences 108, 3163-3168. DOI: 10.1073/pnas.1015975108. <https://doi.org/10.1073/pnas.1015975108> [9] Karen Wynn. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology* 24, 220-251. DOI: 10.1016/0010-0285(92)90008-P. [https://doi.org/10.1016/0010-0285\(92\)90008-P](https://doi.org/10.1016/0010-0285(92)90008-P) [10] Barbara W. Sarnecka; Susan Carey. (2008). How counting represents number: What children must learn and when they learn it. *Cognition* 108, 662-674. DOI: 10.1016/j.cognition.2008.05.007. <https://doi.org/10.1016/j.cognition.2008.05.007> [11] Andrea Bender; Sieghard Beller. (2012). Nature and culture of finger counting: diversity and representational effects of an embodied cognitive tool. *Cognition* 124, 156-182. DOI: 10.1016/j.cognition.2012.05.005. <https://doi.org/10.1016/j.cognition.2012.05.005> [12] Geoffrey B. Saxe. (1981). Body parts as numerals: a developmental analysis of numeration among the Oksapmin in Papua New Guinea. *Child Development* 52, 306-316. DOI: 10.2307/1129244. <https://doi.org/10.2307/1129244> [13] Karenleigh A. Overmann. (2016). The role of materiality in numerical cognition. *Quaternary International* 405, 42-51. DOI: 10.1016/j.quaint.2015.05.026. <https://doi.org/10.1016/j.quaint.2015.05.026> [14] Karenleigh A. Overmann. (2018). Constructing a concept of number. *Journal of Numerical Cognition* 4, 464-493. DOI: 10.5964/jnc.v4i2.161. <https://doi.org/10.5964/jnc.v4i2.161> [15] Hans J. Nissen; Peter Damerow; Robert K. Englund. (1993). *Archaic Bookkeeping: Early Writing and Techniques of Economic Administration in the Ancient Near East*. University of Chicago Press. <https://cdli.earth/publications/1657148> [16] Denise Schmandt-Besserat. (1992). *Before Writing, Volume I: From Counting to Cuneiform*. University of Texas Press. <https://utpress.utexas.edu/9781477325766/> [17] Christopher Woods; Emily Teeter; Geoff Emberling (editors). (2010). *Visible Language: Inventions of Writing in the Ancient Middle East and Beyond*. Oriental Institute Museum Publications 32. <https://isac.uchicago.edu/research/publications/oimp/oimp-32-visible-language-inventions-writing-ancient-middle-east-and> [18] Stephen Chrisomalis. (2010). *Numerical Notation: A Comparative History*. Cambridge University Press. DOI: 10.1017/CBO9780511676062. <https://doi.org/10.1017/CBO9780511676062> [19] Jiajie Zhang; Donald A. Norman. (1994). Representations in distributed cognitive tasks. *Cognitive Science* 18, 87-122. DOI: 10.1207/s15516709cog1801_3. https://doi.org/10.1207/s15516709cog1801_3 [20] Karl Menninger. (1992). *Number Words and Number Symbols: A Cultural History of Numbers*. Dover Publications.

KNO-003 · De las marcas a los símbolos

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

1. Una marca es un acontecimiento antes de ser un signo

La pregunta

¿Qué debe añadirse a una diferencia física para que esa diferencia represente algo para alguien? Una marca es primero un acontecimiento material: se deposita pigmento, se incide una superficie, se rompe una rama o se mueve una mano. Su geometría y su secuencia de fabricación pueden ser observables aunque se desconozca su finalidad. Llamarla signo añade una afirmación funcional: la diferencia se usó o interpretó como algo que mantenía una relación relevante con otra cosa.[18,19] Esa relación puede existir sin comunicación deliberada. Las huellas pueden informar a un rastreador porque fueron causadas por el paso de un animal. Las nubes oscuras pueden orientar la acción porque se correlacionan con el tiempo atmosférico. Tales indicios aportan información a un observador, pero ni el animal ni la nube tienen que haberlos producido para modificar las creencias de otra persona.[21] Una señal suele ser más restringida: una conducta o un artefacto cuya producción participa en la comunicación. Las señales humanas dependen a menudo de que el público reconozca que el productor intenta hacer manifiesto algo. Sin embargo, la producción intencional no garantiza el éxito. Una marca privada cuya regla nadie más puede inferir sigue siendo un intento de señal, no un símbolo público establecido.[16,17] La misma forma puede cambiar de condición según la situación. Una mancha roja puede ser una suciedad, un índice de lesión, una advertencia mostrada deliberadamente o un emblema convencional. La marca física no contiene una categoría semiótica única y permanente. La clasificación depende de la relación, la tarea y la comunidad interpretativa.[9,16,21]

En qué punto estamos

La descripción material va primero. La representación empieza solo cuando la evidencia sostiene una relación entre una forma, un objeto o estado de cosas y un intérprete. Deliberación, capacidad informativa y convencionalidad deben comprobarse por separado, porque una marca hecha a propósito cuya regla nadie más puede inferir sigue siendo una señal intentada y no un símbolo público establecido. La forma física no lleva categoría semiótica permanente.

2. El parecido abre una puerta, pero no fija el significado

La pregunta

Si un signo se parece a su referente, ¿queda ya asegurado su significado? La iconicidad es un parecido percibido o una correspondencia estructural entre un signo y algún aspecto de aquello que denota. Un dibujo puede conservar un contorno; un gesto puede imitar movimiento; una vocalización puede variar en velocidad, duración o tono de maneras que los oyentes relacionan con un acontecimiento. La iconicidad puede facilitar la inferencia de un signo nuevo, sobre todo antes de que exista una convención.[9-11] El parecido es selectivo, no completo. Un boceto de un ave puede destacar las alas, el pico, el vuelo o el tamaño. Qué rasgos cuentan depende del conjunto de

contraste, del medio y de los conocimientos previos del observador. Incluso un icono aparentemente transparente depende, por tanto, de elecciones sobre qué similitud importa.[9] Los experimentos muestran que los participantes pueden crear vocalizaciones no lingüísticas que oyentes sin entrenamiento interpretan por encima del azar, y que la interacción repetida puede convertir formas inicialmente icónicas en signos compartidos. Estos resultados demuestran mecanismos disponibles en condiciones controladas. No muestran que toda palabra hablada comenzara como imitación ni que la iconicidad siga un declive universal.[10,11] Las lenguas combinan iconicidad, arbitrariedad y sistematicidad. Un signo puede parecerse a su referente, ser convencional en su comunidad y formar un patrón con otros signos al mismo tiempo. Estas propiedades son dimensiones, no cajas mutuamente excluyentes, y la convencionalización no tiene por qué borrar la motivación icónica.[9]

En qué punto estamos

La iconicidad puede impulsar la comprensión porque el parecido restringe una primera conjetura. No determina una interpretación única ni elimina la necesidad de atención compartida, retroalimentación y convención aprendida. El parecido es selectivo, así que qué rasgos cuentan ya es una elección, y los experimentos muestran mecanismos disponibles en condiciones controladas, no que toda palabra empezara como imitación.

3. Huellas, señalamiento y anclaje indexical

Una huella causal conecta signo y fuente a través de su producción: el humo procede de la combustión, una pisada del contacto y una cicatriz de un acontecimiento anterior. Un gesto indexical conecta la interpretación con la situación presente: señalar dirige la atención de otra persona hacia un lugar, objeto o acontecimiento. Ninguna de estas relaciones equivale al parecido.[16,17] El señalamiento humano es especialmente potente porque opera dentro de la atención conjunta. Los bebés empiezan a usar y comprender el señalar no solo para pedir objetos, sino también para compartir atención y, en algunos experimentos, para aportar información que otra persona no tiene. El significado del gesto depende de una escena compartida y de inferencias sobre por qué se indica ahora ese objeto.[16,17] La dirección visible de un dedo no determina por completo su mensaje. Señalar una taza puede significar «cógela», «esa es mía», «mira la grieta» o «ponlo ahí». La mirada, el momento, el discurso previo y la tarea común de los participantes reducen las posibilidades. El anclaje indexical reduce la incertidumbre; no elimina la interpretación.[16] Los signos materiales también pueden heredar fuerza indexical de su ubicación. Un marcador junto a un camino, un objeto sellado con una cuenta o un signo fijado a una puerta pueden entenderse en parte por dónde aparecen. Trasladar la misma forma a otro entorno puede cambiar o destruir su función aunque se conserve su figura.[18,19] Esto explica por qué la pérdida de contexto es tan perjudicial en arqueología. Procedencia, asociación, orientación, desgaste y repetición pueden mostrar cómo un objeto marcado participaba en una actividad. Extraer un motivo de esa red y comparar solo su contorno sacrifica gran parte de la evidencia que permitiría distinguir decoración, propiedad, conteo, instrucción o exhibición.[18,19]

En qué punto estamos

Los índices recurren a la historia causal y al contexto presente. Pueden anclar un signo al mundo, pero su interpretación sigue dependiendo de lo que los observadores sepan de la situación y de los propósitos ajenos. La colocación aporta parte de esa fuerza, de modo que la misma forma trasladada a otro entorno puede perder su función con la figura intacta, y por eso la pérdida de procedencia cuesta tanto.

4. Intención y atención compartida

La intención comunicativa es recursiva. Un productor puede actuar para que un público reconozca su intento de dirigir la atención o la comprensión, y el público interpreta el acto en parte al reconocer ese intento. La investigación sobre intencionalidad compartida trata la atención conjunta, la coordinación de papeles y la lectura de motivos como recursos centrales de la cognición cultural humana.[16] Esto no convierte la comunicación en lectura mental sin evidencia. Las personas usan mirada, orientación corporal, marcos convencionales, turnos anteriores y relevancia esperada para inferir qué hace un acto. El malentendido sigue siendo posible y las respuestas públicas revelan si las interpretaciones convergen.[4,16] El desarrollo simbólico infantil muestra que comprender un símbolo implica una relación dual. Una maqueta, una imagen o una palabra son objetos por derecho propio y también representaciones de otra cosa. Los niños pequeños pueden concentrarse tanto en el objeto concreto que la relación representativa se vuelve difícil, mientras que la experiencia con imágenes, lenguaje y ficción favorece un uso cada vez más flexible.[15] Por eso no deben fusionarse comunicación intencional y comprensión simbólica. Alguien puede señalar intencionalmente sin usar un símbolo convencional arbitrario; alguien puede leer con competencia un emblema convencional sin saber quién produjo el ejemplar que tiene delante. La intención concierne a un acto; la convención concierne a una regla social repetible.[15-17]

En qué punto estamos

La atención compartida crea una escena en la que un signo puede interpretarse y corregirse. Ayuda a establecer significado, pero los símbolos estables exigen una historia que va más allá de un acto exitoso. Intención y convención no deben confundirse: la intención concierne a un acto y la convención a una regla social repetible, y una puede darse sin la otra.

5. La interacción convierte conjeturas en convenciones

La semiótica experimental de laboratorio sitúa a participantes en tareas donde se restringe el lenguaje ordinario y deben construirse sistemas nuevos de comunicación. Estos estudios no recrean los orígenes del lenguaje. Exponen, en condiciones controladas, cómo invención, retroalimentación, alineamiento y uso repetido pueden transformar signos.[1-3] En experimentos de comunicación gráfica, las parejas produjeron al principio dibujos relativamente elaborados. Mediante la interacción, las formas exitosas se hicieron más simples y esquemáticas. La mera repetición no bastó: importaron la retroalimentación y la coordinación, y las personas ajenas a la historia de interacción solían tener dificultades para interpretar los signos reducidos.[2] Este patrón muestra cómo la información puede cambiar de lugar. Al principio, gran parte de la pista está visiblemente en el parecido del signo. Más tarde, una marca más breve funciona porque los usuarios recuerdan usos exitosos anteriores. La eficiencia para quienes están dentro puede reducir así la transparencia inmediata para quienes llegan de fuera.[2,3] La investigación de la conversación describe un proceso relacionado de fundamentación compartida. Las expresiones referenciales se refinan conjuntamente hasta que los participantes tienen evidencia suficiente para el propósito actual. Las parejas pueden formar pactos conceptuales y favorecer temporalmente una descripción de un objeto porque pertenece a su historia de interacción compartida.[4,5] Una convención no tiene que negociarse explícitamente ni declararse de forma oficial. Puede emerger de la coordinación repetida cuando una comunidad espera que una forma se use e interprete de una manera particular, observa desviaciones y repara fallos. Lo decisivo no es una asociación privada unánime, sino una práctica pública suficientemente estable.[1,4]

En qué punto estamos

Los símbolos pueden crecer desde la interacción porque la retroalimentación selecciona correspondencias viables y la historia compartida permite comprimir. Su aparente arbitrariedad suele ser el residuo de un camino que quienes llegan después no presenciaron. La mera repetición no produce esa compresión, la coordinación sí, y la eficiencia que ganan los iniciados se paga en transparencia para el recién llegado.

6. La transmisión transforma lo que crea la interacción

Un signo que funciona para una pareja afronta otra prueba: ¿pueden aprenderlo y usarlo personas que no participaron en su invención? La transmisión cultural introduce un cuello de botella. Los aprendices reciben ejemplos limitados, reconstruyen un sistema y transmiten sus propias producciones, no una copia perfecta de cada precedente.[6-8] Los experimentos de aprendizaje iterado muestran que lenguas artificiales pueden volverse más estructuradas y fáciles de aprender a través de cadenas de transmisión. La aprendibilidad por sí sola también puede producir sistemas degenerados que colapsan distinciones. Cuando la transmisión se combina con el uso comunicativo, la presión por comprimir se equilibra con la presión por mantener diferenciados significados relevantes.[6,7] Los experimentos con gestos silenciosos encuentran igualmente contribuciones distintas de interacción y transmisión. Las pantomimas improvisadas pueden hacerse más regulares y sistemáticas cuando aprendices sucesivos también se comunican entre sí. Los cambios resultantes son efectos a escala de población, no evidencia de que ningún participante diseñara de antemano el sistema final.[20] La evolución cultural acumulativa puede conservar modificaciones útiles y permitir que usuarios posteriores construyan sobre ellas. Las tareas de laboratorio muestran que el aprendizaje social puede generar convergencia y complejidad adaptativa, pero el resultado depende de la estructura de la tarea, la retroalimentación disponible y lo transmitido. Estos modelos identifican mecanismos, no un calendario universal de evolución simbólica.[8] La transmisión también crea normatividad. Un aprendiz puede producir una variante comprensible y aun así ser corregido porque la comunidad reconoce una forma habitual. Cuando un precedente puede enseñarse, repetirse y vigilarse, los símbolos adquieren historias de uso correcto e incorrecto que superan la intención individual.[3,6]

En qué punto estamos

La interacción establece un terreno común local; la transmisión comprueba si sobrevive sin sus fundadores. Los sistemas de símbolos se vuelven duraderos cuando siguen siendo a la vez aprendibles y discriminantes bajo reconstrucción repetida, porque la aprendibilidad por sí sola puede colapsar distinciones necesarias. Son resultados de población, no diseño de nadie, e identifican mecanismos y no un calendario.

7. Las lenguas nuevas muestran la emergencia de estructura

Los sistemas de signos domésticos creados por personas sordas sin acceso a una lengua de signos convencional demuestran que el gesto puede desarrollar organización semejante a la lingüística sin un modelo lingüístico completo. Estos sistemas no son simples copias de los gestos de los hablantes de su entorno, aunque surgen en ambientes sociales ricos y continúan moldeados por la interacción.[12] La Lengua de Signos Nicaragüense se desarrolló cuando niños sordos entraron en comunidades donde podían comunicarse con compañeros. La investigación comparativa entre cohortes encontró mayor segmentación y expresión sistemática de acontecimientos de movimiento, lo que muestra que los niños no solo adquirían una lengua terminada, sino que también transformaban una en desarrollo.

[13] La Lengua de Signos Beduina de Al-Sayyid ofrece otro caso comunitario. La investigación documentó organización gramatical sistemática en una lengua relativamente joven. Esta es una trayectoria, no una secuencia simple hacia un punto final predeterminado; diferentes redes sociales, estructuras de cohortes y condiciones de contacto pueden sostener trayectorias diferentes.[14] Estos casos establecen que la estructura lingüística convencional puede emerger dentro de comunidades observables. No proporcionan fósiles transparentes de la primera lengua humana ni implican que las lenguas de signos maduras sean más pictóricas o menos complejas que las habladas. La modalidad cambia los recursos disponibles, no la condición del sistema como lengua.[12-14] La inferencia más sólida es mecanística: el uso repetido entre aprendices que interactúan puede crear contrastes estables y organización combinatoria. El origen histórico de la comunicación simbólica sigue siendo una cuestión más amplia que incluye capacidades, ecologías e historias culturales no reproducidas por ninguna comunidad moderna concreta.[13,14,20]

En qué punto estamos

Las lenguas emergentes permiten observar cómo toman forma la convención y la estructura. Iluminan procesos de formación conservando su propia historia y rechazando el papel de sustitutas vivas de la prehistoria. Redes sociales y estructuras de cohorte distintas sostienen trayectorias distintas, y la modalidad cambia los recursos disponibles para un sistema, no su condición de lengua.

8. Lo que las marcas antiguas pueden y no pueden decirnos

El ocre grabado de la cueva de Blombos demuestra la producción deliberada de patrones geométricos en un contexto seguro de la Edad de Piedra Media. Los análisis microscópicos y tecnológicos pueden abordar cómo se hicieron las líneas y si las acciones formaron una secuencia organizada. El objeto conservado no revela una etiqueta hablada ni un mensaje único.[18] Los estudios arqueológicos más amplios pueden comparar motivos, gestos de fabricación, recurrencia y organización espacial. Las convenciones repetidas entre objetos y contextos refuerzan la afirmación de que las marcas participaban en una práctica compartida. Incluso entonces, «simbólico» puede nombrar proposiciones diferentes, desde una exhibición socialmente reconocible hasta la referencia a una categoría ausente, y cada una requiere su propio apoyo.[18,19] La similitud visual entre épocas o regiones distantes es una evidencia débil de significado idéntico. Cruces, retículas y líneas paralelas simples pueden producirse de forma independiente con facilidad. Una conexión histórica convincente necesita continuidad contextual, reglas de producción compartidas, evidencia de transmisión o prácticas asociadas, no solo parecido.[18,19] La ausencia es igual de difícil de interpretar. Gesto, pintura corporal, fibra, madera, dibujos en arena y habla suelen desaparecer, mientras que piedra y hueso sobreviven de forma selectiva. La marca duradera más antigua no es necesariamente el símbolo más antiguo, y una comunidad sin grafismos conservados pudo poseer comunicación convencional compleja.[19] Por tanto, las afirmaciones deben ascender por una escalera probatoria: marca; fabricación deliberada; recurrencia pautaada; convención transmitida socialmente; función referencial; significado específico recuperado. La arqueología suele sostener las tres primeras y a veces la cuarta. Las dos últimas requieren normalmente un puente inscripcional, etnográfico o independiente de otro tipo.[18,19]

En qué punto estamos

Las marcas antiguas pueden establecer una fabricación sofisticada y una práctica recurrente sin entregar un diccionario. La reconstrucción honesta conserva la distancia entre conducta pauta y significado descifrado. «Simbólico» nombra varias proposiciones distintas, cada una con su propio respaldo, y como los medios duraderos sobreviven de forma selectiva, la marca conservada más antigua no tiene por qué ser el símbolo más antiguo.

9. No hay una sola escalera desde la imagen hasta la palabra

Resulta tentador ordenar huellas, iconos, índices y símbolos arbitrarios como etapas evolutivas sucesivas. Los sistemas reales los mezclan. Un diagrama de señalamiento puede combinar parecido, indexación espacial y notación aprendida; una palabra hablada puede ser históricamente arbitraria y contener simbolismo sonoro; un gesto convencional puede seguir siendo parcialmente icónico.[9] La simplificación tampoco significa siempre progreso hacia la abstracción. Los signos pueden acortarse para usuarios experimentados, hacerse más elaborados cuando el público se diversifica o volverse más icónicos cuando debe introducirse un contraste nuevo. El cambio responde a ecología comunicativa, memoria, medio e instituciones.[2,3,7] La persistencia material modifica la escala de coordinación. Una marca duradera puede inspeccionarse después de que se marche su creador, ser copiada por personas que nunca conocieron al inventor e incorporarse a lugares u objetos que restringen la interpretación. La persistencia posibilita convenciones más amplias, pero también separa los signos de la retroalimentación inmediata que los fundamentó inicialmente.[1,18] El origen de los símbolos no es, por tanto, un momento en el que las marcas adquirieron significado. Es el logro repetido de comunidades que alinean la atención, prueban correspondencias, recuerdan usos exitosos, enseñan a quienes llegan y mantienen distinciones. Las formas importan porque facilitan algunas inferencias y reproducciones; las convenciones importan porque la forma por sí sola nunca fija la relación.[1-3,6]

En qué punto estamos

El paso de las marcas a los símbolos es un cambio en la práctica coordinada, no una secuencia previsible de formas. Parecido, contexto, intención, interacción, memoria y transmisión contribuyen en combinaciones distintas. La simplificación no es progreso hacia la abstracción, y los medios duraderos amplían la coordinación mientras separan los signos de la retroalimentación inmediata que los fundó.

Lo que sabemos y cómo lo sabemos

Los estudios del desarrollo distinguen la identidad física de un objeto de su función representativa y documentan atención conjunta y señalamiento tempranos. Los experimentos de interacción muestran que la retroalimentación y el terreno común pueden estabilizar signos nuevos. Los experimentos de transmisión muestran que la facilidad de aprendizaje y la utilidad comunicativa transforman sistemas entre aprendices. Las lenguas de signos emergentes demuestran cambio estructural en comunidades vivas, mientras que el análisis arqueológico establece fabricación y pautas bajo límites mucho más estrictos respecto al significado.[1-21]

Lo que todavía no sabemos

No conocemos una única secuencia por la que comenzara la comunicación simbólica humana, cuánto dependió la convención temprana de gesto, voz o medios perecederos, ni cuándo marcas particulares empezaron a referirse a entidades ausentes en vez de organizar la acción inmediata. Tampoco tenemos una medida general de cuándo una coordinación local se convierte en norma comunitaria. Los experimentos modernos y las lenguas emergentes revelan mecanismos, pero ningún conjunto de datos conservado observa directamente la primera transición.

Qué podría hacernos cambiar de opinión

Nuevos yacimientos que conservaran objetos marcados repetidos dentro de contextos de actividad resueltos con precisión podrían reforzar las afirmaciones de convención compartida. Evidencia de remontaje, residuos, desgaste o secuencia de producción podría mostrar, en cambio, que marcas aparentemente similares servían para tareas diferentes. Estudios longitudinales más amplios de sistemas nuevos de comunicación podrían separar los efectos de interacción, estructura de red y transmisión. Experimentos capaces de predecir qué convenciones se generalizan a aprendices ajenos a su creación mejorarían el puente entre fundamentación local y simbolismo público.

Si solo recuerdas una cosa

Una marca se convierte en símbolo no por su forma, sino porque una comunidad puede usar, reconocer, transmitir y corregir repetidamente una relación entre esa forma y aquello que representa.

Artículos relacionados

KNO-001 · Una marca en una piedra KNO-002 · Contar antes de los números KNO-004 · De los símbolos a la escritura KNO-005 · Lenguaje oral y transmisión KNO-009 · Archivos, bibliotecas y memoria colectiva

Referencias

[1] Bruno Galantucci. (2005). An Experimental Study of the Emergence of Human Communication Systems. *Cognitive Science* 29(5), 737-767. DOI: 10.1207/s15516709cog0000_34. https://doi.org/10.1207/s15516709cog0000_34 [2] Simon Garrod; Nicolas Fay; John Lee; Jon Oberlander; Tracy MacLeod. (2007). Foundations of Representation: Where Might Graphical Symbol Systems Come From? *Cognitive Science* 31(6), 961-987. DOI: 10.1080/03640210701703659. <https://doi.org/10.1080/03640210701703659> [3] Nicolas Fay; Simon Garrod; Leo Roberts; Nik Swoboda. (2010). The Interactive Evolution of Human Communication Systems. *Cognitive Science* 34(3), 351-386. DOI: 10.1111/j.1551-6709.2009.01090.x. <https://doi.org/10.1111/j.1551-6709.2009.01090.x> [4] Herbert H. Clark; Deanna Wilkes-Gibbs. (1986). Referring as a collaborative process. *Cognition* 22(1), 1-39. DOI: 10.1016/0010-0277(86)90010-7. [https://doi.org/10.1016/0010-0277\(86\)90010-7](https://doi.org/10.1016/0010-0277(86)90010-7) [5] Susan E. Brennan; Herbert H. Clark. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22(6), 1482-1493. DOI: 10.1037/0278-7393.22.6.1482. <https://doi.org/10.1037/0278-7393.22.6.1482> [6] Simon Kirby; Hannah Cornish; Kenny Smith. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105(31), 10681-10686. DOI:

10.1073/pnas.0707835105. <https://doi.org/10.1073/pnas.0707835105> [7] Simon Kirby; Monica Tamariz; Hannah Cornish; Kenny Smith. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141, 87-102. DOI: 10.1016/j.cognition.2015.03.016. <https://doi.org/10.1016/j.cognition.2015.03.016> [8] Christine A. Caldwell; Ailsa E. Millen. (2008). Studying cumulative cultural evolution in the laboratory. *Philosophical Transactions of the Royal Society B* 363(1509), 3529-3539. DOI: 10.1098/rstb.2008.0133. <https://doi.org/10.1098/rstb.2008.0133> [9] Mark Dingemanse; Damián E. Blasi; Gary Lupyan; Morten H. Christiansen; Padraic Monaghan. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences* 19(10), 603-615. DOI: 10.1016/j.tics.2015.07.013. <https://doi.org/10.1016/j.tics.2015.07.013> [10] Marcus Perlman; Rick Dale; Gary Lupyan. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society Open Science* 2(8), 150152. DOI: 10.1098/rsos.150152. <https://doi.org/10.1098/rsos.150152> [11] Marcus Perlman; Gary Lupyan. (2018). People Can Create Iconic Vocalizations to Communicate Various Meanings to Naïve Listeners. *Scientific Reports* 8, 2634. DOI: 10.1038/s41598-018-20961-6. <https://doi.org/10.1038/s41598-018-20961-6> [12] Susan Goldin-Meadow; Heidi Feldman. (1977). The Development of Language-Like Communication Without a Language Model. *Science* 197(4301), 401-403. DOI: 10.1126/science.877567. <https://doi.org/10.1126/science.877567> [13] Ann Senghas; Sotaro Kita; Asli Özyürek. (2004). Children Creating Core Properties of Language: Evidence from an Emerging Sign Language in Nicaragua. *Science* 305(5691), 1779-1782. DOI: 10.1126/science.1100199. <https://doi.org/10.1126/science.1100199> [14] Wendy Sandler; Irit Meir; Carol Padden; Mark Aronoff. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* 102(7), 2661-2665. DOI: 10.1073/pnas.0405448102. <https://doi.org/10.1073/pnas.0405448102> [15] Judy S. DeLoache. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences* 8(2), 66-70. DOI: 10.1016/j.tics.2003.12.004. <https://doi.org/10.1016/j.tics.2003.12.004> [16] Michael Tomasello; Malinda Carpenter; Josep Call; Tanya Behne; Henrike Moll. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28(5), 675-691. DOI: 10.1017/S0140525X05000129. <https://doi.org/10.1017/S0140525X05000129> [17] Ulf Liszkowski; Malinda Carpenter; Tricia Striano; Michael Tomasello. (2006). 12- and 18-Month-Olds Point to Provide Information for Others. *Journal of Cognition and Development* 7(2), 173-187. DOI: 10.1207/s15327647jcd0702_2. https://doi.org/10.1207/s15327647jcd0702_2 [18] Christopher S. Henshilwood; Francesco d'Errico; Royden Yates; Zenobia Jacobs; Chantal Tribolo; Geoff A. T. Duller; Norbert Mercier; Judith C. Sealy; Hélène Valladas; Ian Watts; Ann G. Wintle. (2002). Emergence of Modern Human Behavior: Middle Stone Age Engravings from South Africa. *Science* 295(5558), 1278-1280. DOI: 10.1126/science.1067575. <https://doi.org/10.1126/science.1067575> [19] Francesco d'Errico; Luc Doyon; Ivan Colagè; Alain Queffelec; Emma Le Vraux; Giacomo Giacobini; Bernard Vandermeersch; Bruno Maureille. (2018). From number sense to number symbols: an archaeological perspective. *Philosophical Transactions of the Royal Society B* 373(1740), 20160518. DOI: 10.1098/rstb.2016.0518. <https://doi.org/10.1098/rstb.2016.0518> [20] Yasamin Motamedi; Marieke Schouwstra; Kenny Smith; Jennifer Culbertson; Simon Kirby. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition* 192, 103964. DOI: 10.1016/j.cognition.2019.05.001. <https://doi.org/10.1016/j.cognition.2019.05.001> [21] Terrence W. Deacon. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. W. W. Norton & Company. <https://www.norton.com/books/9780393317541>

KNO-004 · De los símbolos a la escritura

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

Pregunte cuándo empezó la escritura y la respuesta dependerá por completo de qué se esté contando. Huesos con muescas, marcas de alfarero, impresiones de sello, fichas de contabilidad, cuerdas anudadas y esta frase han sido llamadas escritura por alguien. KNO-004 separa seis cosas que esa única palabra funde: marca, símbolo, notación, protoescritura, escritura parcial y escritura plena. Las trata como categorías analíticas y no como peldaños de una escalera, porque el registro muestra sociedades que las cruzan en órdenes distintos, se detienen en puntos distintos y construyen sistemas potentes que nunca fueron escritura ni lo pretendieron. El umbral que usa es la glotografía: codificar un enunciado concreto en una lengua concreta de modo que se recuperen las palabras y no el sentido. Examina cuatro tradiciones llamadas invenciones independientes, un caso aún en disputa genuina y las invenciones documentadas en tiempo histórico, y cierra con lo que la evidencia no resuelve.

En 60 segundos

Una marca puede ser deliberada sin ser un símbolo, y un sistema de símbolos puede ser potente sin codificar lengua. Lo que convierte un sistema de símbolos en escritura, según la definición usada aquí, es la transferencia fonética: el descubrimiento de que un signo puede representar el sonido de su nombre en vez de su significado. Antes de ese paso el sistema solo crece añadiendo signos; después, el inventario se estabiliza y luego se contrae mientras lo expresable deja de tener límite. Asia suroccidental conserva la única secuencia estratificada en la que fichas, envolturas selladas, tablillas numéricas impresas y protocuneiforme pueden examinarse juntos, aunque ese orden registra primera atestación y no descendencia. El corpus del Indo muestra por qué importa el umbral: allí la disputa no versa sobre antigüedad ni sofisticación, sino sobre si los signos codifican lengua, y sigue abierta. Las notaciones numérica, matemática y musical muestran que un sistema cerrado puede ser ilimitado dentro de su dominio y no codificar lengua alguna, no por quedarse corto sino porque nunca lo pretendió. Los silabarios cheroqui y vai muestran que, una vez disponible la idea de escritura, un sistema funcional puede construirse en pocos años.

Por qué importa esta pregunta

La escritura suele tratarse como la frontera entre prehistoria e historia, de modo que qué cuenta como escritura decide de qué sociedades se dice que entraron siquiera en el registro. Leer toda marca repetida como escritura infla el relato más allá de sus objetos; admitir solo escrituras descifradas borra los sistemas de registro que llevaban cuentas y cantidades sin codificar lengua. La distinción encierra además una lección general sobre la evidencia: qué podía hacer un sistema, para qué servía y de dónde venía son tres afirmaciones distintas, y establecer una no establece las otras.

1. El problema con la palabra «escritura»

Preguntar cuándo empezó la escritura obliga primero a decidir qué se cuenta. Marcas en hueso del Paleolítico superior, tallas de cuenta, marcas de alfarero, impresiones de sello, fichas de contabilidad, cuerdas anudadas andinas, notación matemática y esta misma frase han sido llamadas escritura por alguien. No son lo mismo, y las diferencias no son de grado.

La confusión es antigua y tiene causa. La palabra carga dos ideas distintas a la vez. Una es la idea general de comunicación gráfica duradera, hacer una marca que lleve significado a alguien ausente en el espacio o en el tiempo. La otra es la idea específica de codificar lengua, un sistema que captura un enunciado concreto de forma tan completa que quien sabe leerlo puede reproducirlo. Casi toda discusión sobre los orígenes de la escritura resulta ser, al examinarla, una discusión sobre cuál de las dos se está usando.

Este artículo reserva la palabra escritura para el segundo sentido y da al primero su propio vocabulario. Es una estipulación, no un hallazgo, y conviene decirlo abiertamente en vez de colarlo.

En qué punto estamos

Casi toda disputa sobre los orígenes de la escritura es una disputa sobre cuál de dos ideas lleva la palabra: comunicación gráfica duradera, o codificación de un enunciado concreto. Este artículo usa la segunda y da a la primera su propio vocabulario. Es una estipulación y no un hallazgo, por eso se declara sin rodeos, y es lo que hace necesarias las distinciones que siguen.

2. Seis categorías, no seis etapas

Las distinciones que siguen son analíticas. Clasifican lo que un sistema hace, no cuándo apareció ni en qué se convirtió.

Marca. Traza gráfica deliberada cuyo significado, si lo tiene, no se recupera de la traza misma. Una muesca, un trazo, un rayado. Las marcas pueden ser contadores, signos de propiedad, decoración o nada. El problema interpretativo es que una marca y un signo son indistinguibles cuando se pierde el contexto.

Símbolo. Forma gráfica asociada por convención a un referente, siendo la asociación compartida por una comunidad y no privada. Un símbolo lleva significado sin llevar lengua. Un círculo cruzado puede significar un lugar, un dios, un clan o una mercancía, y sus usuarios pueden leerlo hablen la lengua que hablen, **siempre que compartan la convención**. Lo que cruza una frontera lingüística es la convención y no el signo, y un signo cuya convención no se comparte no lleva nada.

Notación. Sistema cerrado de símbolos con reglas de combinación, restringido a un dominio. La notación numérica es el caso más claro: inventario pequeño, sintaxis estricta, productividad ilimitada dentro de su dominio y ninguna capacidad de decir nada fuera de él. La notación no es escritura defectuosa. Es otro instrumento, y puede superar a la escritura en aquello para lo que sirve. La historia comparada de los sistemas numéricos muestra la notación desarrollándose por vías propias y restringidas, en sociedades con escritura y sin ella (Chrisomalis 2010).

Protoescritura. Sistema de símbolos con signos convencionales suficientes, regularidad suficiente y comportamiento combinatorio suficiente para registrar contenido estructurado, pero sin vínculo demostrable con los sonidos ni con la gramática de una lengua concreta. La protoescritura puede registrar qué se debía, a quién y cuánto. No puede registrar por qué, ni qué dijo alguien al respecto.

Escritura parcial. Sistema que sí codifica lengua, pero solo dentro de un dominio acotado y solo para un repertorio restringido de enunciados. Las escrituras administrativas tempranas son el caso tipo: codifican nombres, cantidades y términos de mercancía en una lengua real, con signos fonéticos reales, y nada más. Quien lee recupera las palabras de un recibo y no podría usar el mismo sistema para registrar una queja.

Escritura plena. Sistema capaz de codificar cualquier enunciado de su lengua, incluidos enunciados para los que no fue diseñado. Esta es la propiedad que importa y es la que hace posible el desciframiento en principio, porque un sistema abierto deja huellas gramaticales: palabras funcionales, afijos, concordancia, frecuencias de signo que se comportan como una lengua y no como una lista.

La razón para insistir en seis categorías y no en una línea es que el registro no discurre por la línea. La notación numérica y la escritura plena coexisten y siempre lo han hecho. La protoescritura a veces precede a la escritura en la misma tradición y a veces no tiene descendencia. La escritura parcial a veces se expande hasta ser plena a lo largo de siglos y a veces no se expande en absoluto. Y las categorías son propiedades de sistemas en uso, de modo que unos mismos signos pueden pertenecer a categorías distintas en manos distintas.

En qué punto estamos

Marca, símbolo, notación, protoescritura, escritura parcial y escritura plena clasifican qué hace un sistema, no cuándo apareció ni en qué se convirtió después. Un símbolo lleva significado sin llevar lengua, y lo que cruza una frontera lingüística es la convención y no el signo. La notación no es escritura defectuosa sino otro instrumento, y leer cualquiera de estas categorías como un peldaño convierte una clasificación en una historia.

3. El umbral: la transferencia fonética

Lo que convierte un sistema de símbolos en un sistema de escritura es el descubrimiento de que un signo puede valer por el sonido de su nombre y no por su significado. Cuando el signo de una caña puede usarse para escribir cualquier sílaba que suene como la palabra caña, el inventario deja de ser una lista de cosas y pasa a ser una manera de deletrear. Suele llamarse principio de rebus, y es el mecanismo visible en las transiciones atestiguadas de forma independiente que este artículo examina. **Que toda transición de esa clase en cualquier lugar siguiera esa ruta no se establece aquí**, y los casos que siguen se leen como casos y no como ejemplos de una regla.

Su importancia es estructural, no ingeniosa. Antes de la transferencia fonética el sistema solo puede crecer añadiendo signos y solo puede expresar aquello para lo que tiene signo. Después, el sistema puede expresar cualquier cosa decible, incluidos nombres propios, palabras extranjeras, abstracciones y terminaciones gramaticales, con un inventario que deja de crecer y empieza a encogerse. La firma del cruce no es, por tanto, un signo nuevo sino una economía nueva: inventarios de signos que se estabilizan y luego se contraen mientras el contenido expresable se amplía sin límite.

Esto da un indicador utilizable, y un indicador no es una prueba. Un sistema cuyo número de signos sigue subiendo a medida que se amplía el contenido difícilmente será glotográfico. Un sistema cuyo número de signos baja mientras su contenido se amplía se comporta como se comporta un sistema glotográfico, **lo cual es evidencia y no demostración**: clasificar un sistema exige evidencia convergente, de las frecuencias de signos, de las trazas gramaticales y del rango de contenido efectivamente registrado, y el recuento de signos por sí solo no zanja nada.

En qué punto estamos

La transferencia fonética es el mecanismo que sigue este artículo: un signo capaz de representar el sonido de su nombre deja de ser una entrada de una lista y pasa a ser una forma de deletrear. Su firma es una economía y no una figura, con inventarios que se estabilizan y luego se contraen. Eso da un indicador utilizable y no una prueba.

4. Cuatro tradiciones que suelen tratarse como independientes

Cuatro tradiciones se describen convencionalmente como inventoras de escritura sin modelo. La palabra convencionalmente está haciendo trabajo real en esa frase, y la sección 7 vuelve sobre ello.

4.1 Asia suroccidental

Mesopotamia ofrece el único caso donde algo parecido a la secuencia completa es visible en material estratificado. El cuarto milenio da fichas de arcilla, envolturas selladas, tablillas numéricas impresas y, a finales del cuarto milenio, tablillas protocuneiformes de Uruk, seguidas a lo largo de varios siglos por un cuneiforme capaz de registrar sumerio y después acadio de forma completa. **El orden de esa lista es el de primera atestiguación y no una línea de descendencia.** Los soportes se solapan en el uso, varios circulan a la vez, y cuál de ellos es antecesor de cuál se discute en lugar de estar zanjado.

La propuesta influyente de que las fichas de arcilla son el antecedente directo de los signos, con contadores tridimensionales que se aplanan en impresiones bidimensionales y luego en caracteres dibujados, fue expuesta con detalle por Schmandt Besserat (1992). Nunca ha logrado consenso. Las reseñas cuestionaron tanto el argumento geográfico como la proliferación de formas de ficha de la que dependen las correspondencias de signo (Van De Mieroop 1994). El resumen honesto es que fichas y tablillas tempranas pertenecen a un mismo mundo contable y que una genealogía concreta de forma de ficha a forma de signo sigue en disputa y no establecida.

Trabajos recientes han ampliado el conjunto de antecedentes posibles más allá de las fichas. Un estudio de la iconografía de sellos cilíndricos preliterarios tardíos identifica motivos, entre ellos un paño con flecos y un recipiente reticulado, que reaparecen como signos protocuneiformes usados para las mismas mercancías, y sostiene que iconografía de sello y escritura naciente eran partes de una misma práctica de comunicación sobre arcilla y no una fuente y su derivado (Kelley, Cartolano y Ferrara 2025). Esto no resuelve la cuestión de las fichas. La reencuadra, al retirar el supuesto de que la ascendencia deba ser única.

Dos rasgos del protocuneiforme merecen énfasis porque son para lo que sirven las seis categorías. Sus textos conocidos son administrativos, no muestran sintaxis y no se ha establecido ninguna lectura de ellos como frases en una lengua. Según las definiciones anteriores cae del lado de la protoescritura o, como mucho, del de la escritura parcial, **y esa ubicación es una lectura del corpus conservado y no una propiedad demostrada del sistema.** Varios siglos más lo separan de un cuneiforme que registre sumerio de forma completa.

4.2 Egipto

La tumba U j de Abidos dio unas 160 etiquetas pequeñas de hueso y marfil incisas que el excavador identificó como la escritura egipcia más antigua, de finales del cuarto milenio (Dreyer 1998). Los trabajos de radiocarbono sobre la necrópolis de Umm el Qaab proporcionan el marco cronológico de la

secuencia de tumbas reales del lugar (Görsdorf, Dreyer y Hartung 1998), y los recipientes importados de la tumba han sido reevaluados petrográficamente y contextualmente en fecha más reciente (Ownby, Hartung y Sowada 2022).

El caso egipcio importa por una razón que nada tiene que ver con la prioridad. Las etiquetas más antiguas ya usan signos fonéticos. No hay un aprendizaje largo visible en el material conservado, ninguna etapa atestiguada de símbolos egipcios puramente semánticos adquiriendo despacio valores de sonido. Esa etapa pudo ocurrir sobre materiales que no se conservan; pudo no ocurrir; o pudo ocurrir sobre materiales conservados que no se han reconocido como tales. **Las alternativas no se agotan en las dos primeras**, y la independencia de la invención egipcia sigue sin ser un hecho zanjado en un sentido ni en otro, que es lo que convierte la independencia de la invención egipcia en una pregunta real y no en un hecho asentado.

4.3 China

Las inscripciones oraculares de la capital Shang tardía en Anyang son ya un sistema de escritura maduro cuando aparecen. Codifican una lengua real, con partículas gramaticales y elementos fonéticos productivos, y se pueden leer. La datación por radiocarbono de huesos oraculares sitúa el reinado de Wu Ding, del que procede el primer corpus grande, hacia el siglo XIII a.C., con el final Shang hacia mediados del XI (Liu et al. 2021).

El problema es el hueco que tienen detrás. Se conocen signos incisos neolíticos en yacimientos mucho más antiguos, entre ellos marcas sobre caparazón de tortuga y hueso en Jiahu, del séptimo milenio a.C., publicadas bajo un signo de interrogación y no como afirmación (Li et al. 2003). Nada demuestra continuidad entre esas marcas y la escritura Shang, y las separan varios miles de años. El caso chino es por tanto la ilustración más clara de un problema general: una escritura madura sin ascendencia visible nos dice que la ascendencia se escribió sobre algo que se pudre.

4.4 Mesoamérica

Mesoamérica aporta el caso más fuerte de invención genuinamente independiente entre los aquí tratados, porque no se ha propuesto ninguna ruta de contacto plausible con las escrituras del Viejo Mundo. **Es un argumento desde la ausencia de una ruta propuesta**, que es la forma que la sección 8 advierte como la más débil disponible, y así se ofrece. Aporta también la disputa más instructiva sobre evidencia.

El Bloque de Cascajal, una losa de serpentina con sesenta y dos signos, se publicó como la escritura más antigua del Nuevo Mundo y se asignó a la fase San Lorenzo, que termina hacia el 900 a.C. (Rodríguez Martínez et al. 2006). El bloque no fue excavado. Unos trabajadores de carretera lo recuperaron de un montón de escombros a finales de los años noventa, y la publicación dice con claridad que está desligado de su contexto estratigráfico y datado por asociación con el material circundante y por análisis iconográfico. Sea lo que sea el bloque, su fecha descansa en una inferencia sobre un hallazgo no estratificado, y eso debería decirse cada vez que se cita.

Las piedras labradas zapotecas del valle de Oaxaca ofrecen un caso temprano mejor anclado, con el Monumento 3 de San José Mogote que lleva un nombre de día calendárico junto a una figura de cautivo (Marcus 1976). La escritura maya es plenamente glotográfica y descifrable, y un signo de día recuperado en contextos sellados de San Bartolo retrasa la escritura calendárica maya atestiguada hasta aproximadamente los siglos III a II a.C. (Stuart et al. 2022), tras la publicación anterior de texto jeroglífico del mismo yacimiento (Saturno, Stuart y Beltrán 2006).

En qué punto estamos

Las cuatro tradiciones se llaman convencionalmente invenciones independientes, y el adverbio carga con el peso. La secuencia mesopotámica es un orden de primera atestación y no una línea de descendencia, la genealogía de ficha a signo nunca alcanzó consenso, y el Bloque de Cascajal se data por asociación desde un hallazgo sin estratigrafía. La independencia mesoamericana descansa en que no se ha propuesto ninguna ruta de contacto.

5. El caso que muestra por qué importa el umbral

Los signos del valle del Indo son la mejor prueba disponible del marco, precisamente porque su estatus está sin resolver.

Una posición sostiene que el corpus no es escritura en absoluto. Las inscripciones son extremadamente cortas, con una media inferior a cinco signos y la más larga conocida en diecisiete; una cuarta parte del inventario de signos aparece una sola vez; los patrones de repetición de signos no se comportan como en las escrituras fonéticas; y faltan los correlatos arqueológicos de la producción de manuscritos. En esta lectura los signos son un sistema emblemático no lingüístico, comparable a las convenciones de sello y divisa del Próximo Oriente (Farmer, Sproat y Witzel 2004).

La posición contraria aplica medidas estadísticas de estructura de secuencia. La entropía condicional calculada sobre el corpus del Indo queda más cerca de la de lenguas conocidas que de la de sistemas de signos no lingüísticos, lo que los autores ofrecen como evidencia, explícitamente no como prueba, de que los signos codifican lengua (Rao et al. 2009). El método fue después discutido en su propio terreno, con una batería más amplia de pruebas estadísticas que, según se informa, no separa de forma fiable los sistemas lingüísticos de los no lingüísticos (Sproat 2014).

De ello se siguen dos cosas. Primera, la disputa no versa sobre la antigüedad de los signos del Indo ni sobre la sofisticación de aquella sociedad. Versa sobre si un sistema simbólico cruzó el umbral glotográfico, que es exactamente la distinción que establece la sección 2. Segunda, la estructura estadística por sí sola no lo ha resuelto hasta ahora. Los textos cortos llevan poca señal, y el argumento desde la estadística de secuencias ha demostrado ser reversible.

En qué punto estamos

El corpus del Indo es la mejor prueba disponible del marco precisamente porque su estatus no está resuelto. Inscripciones cortas, un inventario con muchos signos de aparición única y la ausencia de correlatos de producción manuscrita sostienen una lectura no lingüística; la entropía condicional lo sitúa más cerca de lenguas conocidas; una batería más amplia de pruebas no separó ambas clases con fiabilidad.

6. Sistemas que nunca pretendieron ser escritura

Tratar la escritura como destino hace que cualquier otro sistema de registro parezca un intento que se quedó corto. Es un error historiográfico serio y distorsiona para qué servían esos sistemas.

La notación numérica es el contraejemplo más simple. Es cerrada, precisa, ilimitada dentro de su dominio y del todo incapaz de codificar lengua, y nunca ha estado bajo presión para cambiar (Chrisomalis 2010). Las notaciones matemática, química y musical se comportan igual. Son semasiográficas por diseño: representan contenido directamente y no a través de la lengua, que es la razón de que crucen fronteras lingüísticas sin traducción.

Los sistemas de registro mesoamericanos y andinos se han analizado en esos mismos términos, como literacidades por derecho propio y no como aproximaciones a la escritura alfabética (Boone y Mignolo 1994). El quipu inca, un sistema de cuerdas anudadas, es el caso más difícil de ese grupo. Su función contable no está en duda; la propuesta de que codifica además contenido no numérico mediante una codificación binaria combinatoria es una hipótesis viva y no un hecho demostrado (Urton 2003).

El punto general es que la semasiografía no es una etapa. Es una alternativa permanente, coexiste con la escritura allí donde hay escritura y en varios dominios es la mejor herramienta.

En qué punto estamos

Las notaciones numérica, matemática, química y musical son cerradas, precisas, ilimitadas dentro de su dominio e incapaces de codificar lengua, y nunca han estado bajo presión para cambiar. Los sistemas mesoamericanos y andinos se han analizado como literacidades por derecho propio y no como aproximaciones. La función contable del quipu no está en duda mientras su codificación no numérica sigue siendo una hipótesis.

7. Con qué rapidez puede darse el paso: las invenciones documentadas

La evidencia más fuerte contra una escalera universal lenta procede de casos observados en tiempo histórico.

El silabario cherokee fue ideado por Sequoyah y estaba en uso a principios de la década de 1820, con reconocimiento del consejo cherokee en 1821 (Cushman 2011). Es un silabario completo para una lengua real, producido sin una secuencia larga de desarrollo, por alguien que sabía que la escritura existía entre sus vecinos. Se ha aportado trabajo arqueológico a la cuestión de su independencia (Weeks y Tankersley 2011).

El silabario vai de Liberia siguió en torno a una década después, y la cuestión de si recibió influencia cherokee a través de colonos liberianos se ha examinado en detalle en vez de darse por supuesta en un sentido o en otro (Tuchscherer y Hair 2002).

Estos casos aíslan la variable que los casos arqueológicos no pueden aislar. Lo que lleva siglos no es la invención de la escritura. Es la invención de la idea de que la lengua puede escribirse. En los dos casos documentados en tiempo histórico, una vez disponible esa idea, incluso de segunda mano y sin conocimiento de ninguna escritura concreta, se construyó un sistema funcional en unos pocos años.

Dos casos es lo que esto respalda, y muestran que el paso puede ser rápido, no que lo sea siempre. Esto es difusión de estímulo, y significa que las afirmaciones de independencia para los casos antiguos exigen un argumento sobre qué podían conocer sus inventores, y no solo una demostración de que sus signos no se parecen a los de nadie.

En qué punto estamos

El silabario cheroqui estaba en uso a principios de la década de 1820 y recibió reconocimiento del consejo en 1821, y el vai siguió una década después, con su posible influencia cheroqui examinada y no supuesta. Estos casos aíslan lo que los arqueológicos no pueden: lo lento no es construir una escritura sino que llegar la idea de que la lengua puede escribirse.

8. Lo que la evidencia no puede resolver

Cuatro límites merecen enunciarse en vez de gestionarse.

Conservación. La escritura temprana se conserva sobre arcilla, piedra, hueso y concha. También se escribió sobre papiro, corteza, bambú, madera, tela y piel, que no se conservan en la mayoría de los climas. Una aparición aparentemente súbita de una escritura madura es compatible con un desarrollo largo e invisible, y excavar más en un soporte que no conserva no puede descartarlo. **Compatible no es lo mismo que probable**, y el sesgo corre en un solo sentido: puede ocultar un desarrollo que ocurrió y no puede fabricar uno que no ocurrió.

Afirmaciones negativas sobre independencia. Mostrar que dos escrituras difieren no muestra que una se inventara sin conocimiento de la otra, porque la difusión de estímulo transmite la idea y no las formas. La independencia es por tanto un argumento sobre oportunidad y contacto, y esos argumentos se debilitan según se acumula evidencia de contacto a larga distancia.

Corpus sin descifrar. Para una escritura sin bilingüe, sin lengua descendiente conocida y con textos cortos, la pregunta de si es escritura puede no tener respuesta con los métodos actuales. Es una afirmación sobre método, no un veredicto sobre la sociedad.

Las categorías mismas. Los seis términos de la sección 2 son instrumentos, no clases naturales. Una partición distinta produciría casos frontera distintos. Lo que la partición no debe hacer es convertirse en cronología, y la afirmación principal de este artículo es que la lectura cronológica habitual de esos términos es el error que hay que evitar.

En qué punto estamos

Cuatro límites se declaran en vez de gestionarse. La conservación puede ocultar un desarrollo que ocurrió y no puede fabricar uno que no ocurrió. Mostrar que dos escrituras difieren no muestra invención independiente, porque la difusión de estímulo transmite la idea y no las formas. Un corpus sin descifrar puede no ser resoluble con los métodos actuales. Y los seis términos son instrumentos, no clases naturales.

9. Conclusión

La escritura no surgió una sola vez ni por una sola vía. Los sistemas de símbolos se hicieron escritura cuando los signos empezaron a llevar sonido además de sentido, y ese paso se dio en momentos distintos, por razones distintas, en tradiciones que llegaron a él desde puntos de partida distintos. Algunas lo dieron pronto y se expandieron despacio. Una parece haberlo dado de inmediato. Otras construyeron sistemas de registro de gran potencia y nunca lo dieron, porque sus fines no lo requerían.

La pregunta útil no es cuándo empezó la escritura. Es qué podía hacer un sistema dado, para quién, y qué no podía hacer. Responderla exige las distinciones, y prohíbe la escalera.

En qué punto estamos

La escritura no surgió una sola vez ni por una sola ruta. Los signos empezaron a llevar sonido además de sentido en momentos distintos y por razones distintas, en tradiciones que llegaron ahí desde puntos de partida distintos, mientras otras construían sistemas de registro potentes y nunca dieron el paso, porque sus fines no lo requerían. La pregunta útil es qué podía hacer un sistema y qué no.

Lo que sabemos y cómo lo sabemos

El material estratificado del cuarto milenio en Asia suroccidental conserva la única secuencia en la que fichas, envolturas selladas, tablillas numéricas impresas y protocuneiforme pueden examinarse juntos, aunque ese orden es de primera atestación y no de descendencia. Las invenciones modernas documentadas fijan el ritmo: el silabario cheroqui estaba en uso a principios de la década de 1820 y recibió reconocimiento del consejo en 1821, y el silabario vai siguió en torno a una década después. Las notaciones numérica, matemática, química y musical muestran que un sistema cerrado puede ser ilimitado dentro de su dominio y no codificar lengua alguna.

Lo que todavía no sabemos

Si los signos del Indo codifican lengua sigue sin resolverse: las medidas de entropía los situaron más cerca de lenguas conocidas, ese método fue discutido en su propio terreno y no existe ningún bilingüe. Si las fichas de arcilla son antecesoras directas de los signos cuneiformes nunca ha alcanzado consenso. Si toda transición a la escritura siguió la vía de la transferencia fonética no queda establecido aquí. Si el khipu inka codifica contenido no numérico sigue siendo una hipótesis viva y no un hecho demostrado.

Qué podría hacernos cambiar de opinión

Un bilingüe, una lengua descendiente conocida o textos bastante más largos llevarían la cuestión del Indo del método a la evidencia. Excavar en un medio que conserve podría revelar un desarrollo largo tras una escritura de apariencia súbita, aunque ese sesgo actúa en una sola dirección y puede ocultar un desarrollo que ocurrió, no fabricar uno que no ocurrió. La acumulación de evidencia de contacto a larga distancia debilitaría las afirmaciones de independencia, que descansan en la oportunidad y no en una demostración.

Si solo recuerdas una cosa

La escritura empieza cuando un signo puede llevar el sonido de su nombre en vez de su significado; los demás sistemas de registro son alternativas y no intentos fallidos.

Artículos relacionados

- KNO-002 · *Contar antes de los números*
- KNO-003 · *De las marcas a los símbolos*
- KNO-005 · *Lengua oral y transmisión*

Referencias

Boone, Elizabeth Hill, y Walter D. Mignolo (eds.). 1994. *Writing Without Words: Alternative Literacies in Mesoamerica and the Andes*. Durham: Duke University Press. ISBN 978-0-8223-1377-9 (tapa dura); 978-0-8223-1388-5 (rústica). Chrisomalis, Stephen. 2010. *Numerical Notation: A Comparative History*. Cambridge: Cambridge University Press. ISBN 978-0-521-87818-0. Cushman, Ellen. 2011. «The Cherokee Syllabary: A Writing System In Its Own Right». *Written Communication* 28 (3): 255–281. DOI 10.1177/0741088311410172. Daniels, Peter T., y William Bright (eds.). 1996. *The World's Writing*

Systems. New York y Oxford: Oxford University Press. ISBN 978-0-19-507993-7. Dreyer, Günter. 1998. *Umm el-Qaab I: Das prädynastische Königsgrab U-j und seine frühen Schriftzeugnisse*. Deutsches Archäologisches Institut, Abteilung Kairo, Archäologische Veröffentlichungen 86. Mainz: Verlag Philipp von Zabern. ISBN 978-3-8053-2486-1. Farmer, Steve, Richard Sproat, y Michael Witzel. 2004. «The Collapse of the Indus-Script Thesis: The Myth of a Literate Harappan Civilization». *Electronic Journal of Vedic Studies* 11 (2): 19–57. URL estable: <https://safarmer.com/fsw2.pdf> Ferrara, Silvia, Barbara Montecchi, y Miguel Valério (eds.). 2024. *Writing from Invention to Decipherment*. Oxford: Oxford University Press. ISBN 978-0-19-890874-6. DOI 10.1093/oso/9780198908746.001.0001. Görsdorf, Jochen, Günter Dreyer, y Ulrich Hartung. 1998. «New 14C Dating of the Archaic Royal Necropolis Umm el-Qaab at Abydos (Egypt)». *Radiocarbon* 40 (2): 641–647. DOI 10.1017/S0033822200018579. Houston, Stephen D. (ed.). 2004. *The First Writing: Script Invention as History and Process*. Cambridge: Cambridge University Press. ISBN 978-0-521-83861-0. Kelley, Kathryn, Mattia Cartolano, y Silvia Ferrara. 2025. «Seals and signs: tracing the origins of writing in ancient South-west Asia». *Antiquity* 99 (403): 64–82. DOI 10.15184/aqy.2024.165. Li, Xueqin, Garman Harbottle, Juzhong Zhang, y Changsui Wang. 2003. «The earliest writing? Sign use in the seventh millennium BC at Jiahu, Henan Province, China». *Antiquity* 77 (295): 31–44. DOI 10.1017/S0003598X00061329. Liu, Kexin, Xiaohong Wu, Zhiyu Guo, Sixun Yuan, Xingfang Ding, Dongpo Fu, y Yan Pan. 2021. «Radiocarbon Dating of Oracle Bones of Late Shang Period in Ancient China». *Radiocarbon* 63 (1): 155–175. DOI 10.1017/RDC.2020.90. Marcus, Joyce. 1976. «The Origins of Mesoamerican Writing». *Annual Review of Anthropology* 5: 35–67. DOI 10.1146/annurev.an.05.100176.000343. Ownby, Mary F., Ulrich Hartung, y Karin Sowada. 2022. «The jars from Tomb U-j – a further petrographic and contextual reassessment». *Ägypten und Levante* 32: 231–268. DOI 10.1553/AEundL32s231. Rao, Rajesh P. N., Nisha Yadav, Mayank N. Vahia, Hrishikesh Joglekar, R. Adhikari, e Iravatham Mahadevan. 2009. «Entropic Evidence for Linguistic Structure in the Indus Script». *Science* 324: 1165. DOI 10.1126/science.1170391. Rodríguez Martínez, Ma. del Carmen, Ponciano Ortíz Ceballos, Michael D. Coe, Richard A. Diehl, Stephen D. Houston, Karl A. Taube, y Alfredo Delgado Calderón. 2006. «Oldest Writing in the New World». *Science* 313: 1610–1614. DOI 10.1126/science.1131492. Saturno, William A., David Stuart, y Boris Beltrán. 2006. «Early Maya Writing at San Bartolo, Guatemala». *Science* 311 (5765): 1281–1283. DOI 10.1126/science.1121745. Schmandt-Besserat, Denise. 1992. *Before Writing, Vol. I: From Counting to Cuneiform*. Austin: University of Texas Press. ISBN 978-0-292-70783-2. Sproat, Richard. 2014. «A statistical comparison of written language and nonlinguistic symbol systems». *Language* 90 (2): 457–481. DOI 10.1353/lan.2014.0031. Stuart, David, Heather Hurst, Boris Beltrán, y William Saturno. 2022. «An early Maya calendar record from San Bartolo, Guatemala». *Science Advances* 8 (15): eabl9290. DOI 10.1126/sciadv.abl9290. Tuchscherer, Konrad, y P. E. H. Hair. 2002. «Cherokee and West Africa: Examining the Origins of the Vai Script». *History in Africa* 29: 427–486. DOI 10.2307/3172173. Urton, Gary. 2003. *Signs of the Inka Khipu: Binary Coding in the Andean Knotted-String Records*. Austin: University of Texas Press. ISBN 978-0-292-78540-3. Van De Mieroop, Marc. 1994. Reseña de *Before Writing*, vol. I, de Denise Schmandt-Besserat. *The Classical World* 87 (6): 501. Weeks, Rex, y Kenneth Tankersley. 2011. «Talking leaves and rocks that teach: the archaeological discovery of Sequoyah's oldest written record». *Antiquity* 85 (329): 978–993. DOI 10.1017/S0003598X00068435.

KNO-005 · Lenguaje oral y transmisión

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

Pregunte cómo se recordaba antes de la escritura y la pregunta cuela tres supuestos: que recordar era la dificultad, que la escritura fue la solución, que la transmisión oral es una sola cosa. KNO-005 formula otra más estrecha: dado un cuerpo de conocimiento, una comunidad y ningún soporte gráfico, qué determina cuánto sobrevive a una generación y a qué coste. Distingue seis modos cuya fidelidad difiere mucho, del relato casual a la recitación con verificación institucional, tratados como categorías y no etapas. Examina los dos polos: el cantor que compone en vez de memorizar, y la comunidad que recita en órdenes permutados para que la corrupción se vuelva audible. Pondera qué transportan demostrablemente las cadenas largas y con qué puede contrastarse una afirmación de tiempo profundo, y revisa la gran divisoria. Cierra con la asimetría: una tradición que nunca se registró no deja nada que excavar.

En 60 segundos

La transmisión oral no es una práctica sino una familia de técnicas, y no son tecnologías comparables. El relato casual se degrada deprisa, porque nada en la práctica se opone a ello. La recitación verificada es el régimen de fidelidad oral más fuerte conocido. La diferencia la produce la restricción, no el esfuerzo. Un verso en una tradición fuerte está sostenido a la vez por la métrica, la rima o la aliteración, la melodía, la imagería de su escena, el fraseo formulario de su hueco métrico y la secuencia narrativa. Cada una deja muchas opciones; todas juntas dejan a menudo una, y quien no recuerda un verso regenera uno que encaja. Se siguen tres consecuencias: la fidelidad es específica de dominio, la prosa sin restricción es el caso frágil, y fidelidad y creatividad son ajustes distintos, no opuestos. En un polo, el cantor eslavo meridional construye una ejecución distinta cada vez, con idéntica trama y otras palabras, y no lo vive como un fallo. En el otro, los especialistas recitan un texto en órdenes sistemáticamente permutados, y una corrupción queda expuesta por el desacuerdo entre las formas: redundancia usada para detección de errores, sin teoría de la información.

Por qué importa esta pregunta

Un argumento influyente de mediados del siglo XX sostuvo que la escritura alfabética reestructura el pensamiento y habilita lógica, historia y escepticismo que la oralidad no sostiene. El trabajo comparativo sustituyó ese modelo autónomo por otro en el que lo que la escritura hace depende de las prácticas en que está inserta, y de la tecnología sola no se sigue consecuencia general. Tratar la transmisión oral como etapa halaga a la escritura, al acreditarle la solución de un problema que los regímenes orales fuertes ya habían resuelto, y oscurece lo que sí cambió: coste, portabilidad, acceso aleatorio y registro de lo arbitrario. Y cada pérdida se lleva cuanto se conservara solo en una lengua y solo en la memoria.

1. Una pregunta mal planteada

Cómo recordaba la gente antes de la escritura es una pregunta que tiene respuesta, y la respuesta engaña, porque la pregunta cuela tres supuestos. Supone que recordar era la dificultad. Supone que la escritura fue la solución a esa dificultad. Y supone que la transmisión oral es una sola cosa sobre la que cabe una afirmación general.

Ninguno de los tres resiste el examen. Las sociedades orales no luchaban principalmente por retener conocimiento; muchas retenían cuerpos concretos de conocimiento con fidelidad alta, por los mecanismos que describe la sección 3. Los usos más antiguos de la escritura que se conservan son en buena parte administrativos y no mnemotécnicos, lo cual es una afirmación sobre lo que se conserva y no un motivo único de su adopción. Y las prácticas agrupadas bajo tradición oral difieren entre sí más de lo que cualquiera de ellas difiere de la escritura.

La pregunta mejor es más estrecha y tiene respuesta. Dado un cuerpo de conocimiento, una comunidad y ningún soporte gráfico, qué mecanismos determinan cuánto sobrevive a una generación, y qué cuestan esos mecanismos.

En qué punto estamos

La pregunta habitual supone que recordar era la dificultad, que la escritura la resolvió y que la transmisión oral es una sola cosa. Ninguno de los tres supuestos resiste. Los usos más antiguos que se conservan dicen algo sobre lo que se conserva, no sobre el motivo de su adopción. Otra pregunta más estrecha la sustituye, y responderla empieza por no tratar la transmisión oral como una sola práctica.

2. Seis modos, no uno

Como en KNO-004, las divisiones siguientes clasifican lo que una práctica hace. No son etapas y no forman secuencia.

Relato casual. Alguien cuenta a otro lo que pasó. La fidelidad tiende a ser baja y a degradarse deprisa, porque nada en la práctica está dispuesto para oponerse a ello; donde un relato casual se repite bajo restricción, se comporta de otro modo. Es el modo que suelen estudiar los experimentos psicológicos de cadenas de transmisión, y es el menos parecido a las tradiciones que llevan conocimiento a lo largo de siglos.

Tradición narrativa. Relatos transmitidos entre generaciones con esqueleto argumental estable y superficie inestable. Lo que persiste es estructura, función de personaje y secuencia. La literalidad no persiste ni se espera que lo haga. El trabajo filogenético comparado sobre corpus de cuentos populares trata exactamente esa señal, usando el esqueleto estable para reconstruir descendencia, y encuentra linaje recuperable a lo largo de cadenas muy largas (Graça da Silva y Tehrani 2016).

Composición formularia en ejecución. El modo descrito por Parry y Lord para la épica eslava meridional y aplicado al material homérico. El cantor no reproduce un texto. Lo construye, en vivo, a partir de un repertorio de fórmulas con forma métrica y de temas recurrentes, y construye otro distinto en la ocasión siguiente (Lord 2019). La misma canción es, según esto, un tipo y no un texto, y preguntar si una ejecución es fiel es hacer a la práctica la pregunta equivocada.

Transmisión poética restringida. Verso mantenido estable por restricciones simultáneas: métrica, rima, aliteración, melodía, imaginería y fraseo formulario actuando a la vez. Las restricciones no ayudan al cantor a esforzarse más. Recortan el espacio de lo que puede ir en un hueco hasta que caben muy pocos candidatos, y por eso el recuerdo parcial reconstruye en vez de fallar (Rubin 1995).

Transmisión ritual literal. Habla cuyo valor depende de la literalidad exacta, transmitida con el objetivo explícito de cambio cero. Aquí la fidelidad es una exigencia social sostenida por adiestramiento y por sanción, no una propiedad del material.

Transmisión con verificación institucional. Transmisión literal con un aparato añadido para detectar corrupción, enseñado a especialistas durante años, en el que el texto se recita en órdenes deliberadamente alterados de modo que cualquier error se vuelve audible. Es el régimen de fidelidad oral más fuerte que se conoce y se trata en la sección 4.

La razón para las seis categorías es de nuevo que una sola palabra oculta las diferencias que importan. Relato casual y recitación verificada son ambos orales, y no son tecnologías comparables.

En qué punto estamos

Los seis modos se separan por lo que hace una práctica, no por su lugar en una secuencia: relato casual, tradición narrativa, composición en ejecución, verso restringido, habla ritual literal y recitación con verificación institucional. Las divisiones son clasificaciones y no etapas, y los dos extremos son orales sin ser tecnologías comparables. Lo que todavía no explican es por qué los casos fuertes se sostienen.

3. La fidelidad se construye con restricción, no con esfuerzo

El hallazgo central del trabajo cognitivo sobre tradiciones orales es que la estabilidad procede de la acción simultánea de múltiples restricciones, cada una débil por separado.

Un verso en una tradición fuerte está restringido por su métrica, por su posición en un esquema de rima o aliteración, por la línea melódica con que se canta, por la imaginería de la escena a la que pertenece, por el fraseo formulario disponible para ese hueco métrico y por la secuencia narrativa. Cualquiera de ellas por separado deja muchas opciones. Todas juntas dejan a menudo una. Quien ejecuta y no recuerda un verso no se queda en blanco; las restricciones regeneran un verso que encaja, y el verso regenerado es con frecuencia el original o algo muy próximo (Rubin 1995).

De ahí se siguen tres consecuencias, y explican casi todas las paradojas aparentes del campo.

Primera, **la fidelidad es específica de dominio.** Una misma comunidad puede transmitir una genealogía con exactitud casi perfecta y una instrucción práctica mal, porque una está inserta en una estructura de restricción y la otra no.

Segunda, **la prosa sin restricción es el caso frágil.** Carece de casi toda la maquinaria anterior, y por eso las tradiciones que deben conservarse tienden a versificarse, a cantarse o a ligarse a fórmulas fijas aunque su contenido no sea poético. La prosa sostenida por otras restricciones, como un orden litúrgico fijo, no es frágil del mismo modo.

Tercera, **fidelidad y creatividad no son opuestos sino ajustes distintos.** La composición en ejecución sacrifica estabilidad literal y compra productividad ilimitada. La recitación verificada sacrifica productividad y compra estabilidad casi literal. Las dos son orales, y una sociedad puede operar ambas a la vez para cuerpos de contenido distintos.

En qué punto estamos

La estabilidad procede de muchas restricciones individualmente débiles que actúan a la vez y juntas dejan a menudo una opción donde cada una deja muchas, y quien no recuerda un verso regenera uno que encaja. El verso regenerado es a menudo el original, no siempre. La fidelidad es específica de dominio, la prosa sin restricción es el caso frágil, y los casos más claros de cada ajuste son los polos.

4. Los dos polos

4.1 El cantor que no memoriza

El trabajo de campo de Parry y Lord estableció que el cantor épico eslavo meridional, al que se pide repetir una canción, produce una ejecución de longitud similar y argumento idéntico con una literalidad sustancialmente distinta, y no lo vive como un fallo. No está recitando; está componiendo dentro de un sistema (Lord 2019).

El resultado reencuadra muchas cosas. Significa que la pregunta de si una tradición oral conserva un texto original puede estar vacía para géneros enteros, porque no hubo texto original. Significa también que la estabilidad en el plano de trama, personaje y tema es compatible con la inestabilidad en el plano de las palabras, de modo que una tradición puede ser muy conservadora sobre qué ocurre e indiferente sobre cómo se dice.

4.2 La comunidad que construye una suma de verificación

En el polo opuesto están los regímenes de transmisión diseñados para hacer el error detectable y no meramente improbable. La tradición de recitación védica del sur de Asia es el ejemplo estándar. Junto a la recitación continua, los especialistas aprenden a recitar el mismo texto en órdenes sistemáticamente permutados, palabra a palabra en aislamiento, luego en pares solapados y después en patrones entrelazados más elaborados. Como una corrupción altera lo que producen las formas permutadas, el desacuerdo entre las formas de recitación la deja al descubierto.

Conviene enunciar el punto estructural en términos generales. Esto es redundancia usada para detección de errores: el mismo contenido codificado de varias maneras, de modo que un fallo en una codificación resulta visible frente a las otras. Es el mismo principio que subyace a las sumas de verificación en transmisión de datos, alcanzado sin ninguna teoría de la información y mantenido por adiestramiento institucional durante un periodo muy largo.

La cautela que acompaña a este ejemplo pertenece al texto y no solo a una nota. Este artículo describe la técnica porque está bien atestiguada en la literatura especializada sobre recitación en el sur de Asia; las fuentes concretas que figuran en la bibliografía han sido verificadas como publicaciones, pero **no se ha confirmado directamente que traten estas formas concretas de recitación**, y así consta en la lista de incertidumbres.

En qué punto estamos

En un polo el cantor compone dentro de un sistema y produce la misma trama con otras palabras sin vivirlo como un fallo. En el otro, la recitación en órdenes permutados convierte una corrupción en desacuerdo entre formas, que es redundancia usada para detectar errores. Las fuentes listadas están verificadas como publicaciones, no como tratamiento de esas formas. Qué llevan esas cadenas en plazos muy largos es otra cuestión.

5. Profundidad: qué pueden llevar las cadenas largas

Dos líneas de evidencia informan sobre cuánto sobrevive a una transmisión oral muy larga, y sostienen grados de confianza distintos.

Linaje reconstruible. La comparación filogenética de corpus de cuentos trata los relatos como rasgos heredables y reconstruye su descendencia a partir de su distribución entre lenguas, recuperando una señal que los autores interpretan como de profundidad protoindoeuropea (Graça da Silva y Tehrani 2016). **El resultado depende del modelo:** se sigue del método filogenético y de la codificación de relatos que esos autores adoptan, y en esa literatura se defienden otras codificaciones y otros modelos. Lo que respalda es que algunos rasgos estructurales sobreviven a cadenas extremadamente largas. No establece que un contenido concreto de un relato concreto sea antiguo.

Correspondencia con acontecimientos datados de forma independiente. Una afirmación más fuerte y más discutida es que ciertas tradiciones orales codifican información sobre acontecimientos datables por otros medios. Tradiciones aborígenes australianas que describen la inundación de tierras costeras se han comparado con la curva postglacial del nivel del mar y se ha sostenido que conservan conocimiento de líneas de costa anegadas hace más de siete mil años (Nunn y Reid 2016). **Es una hipótesis publicada y aquí se recoge como tal**, no como una correspondencia establecida.

El problema metodológico de las afirmaciones de este tipo es general y quedó expuesto antes que este caso: las afirmaciones orales de tiempo profundo son estructuralmente difíciles de refutar y de creer, porque la evidencia que las confirmaría y la que las refutaría son a menudo la misma evidencia leída de otro modo (Henige 2009). No se ha localizado, durante la preparación de este artículo, ninguna refutación publicada que aborde específicamente el argumento costero australiano, lo cual es una afirmación sobre esta búsqueda y no un respaldo.

Trabajos recientes extienden el enfoque al vulcanismo y a rasgos geológicos, leyendo de nuevo tradiciones orales contra acontecimientos datados de forma independiente (Franks, Nunn y McCallum 2024; Franks et al. 2025). El campo es pequeño y su autoría se solapa, lo cual es relevante al ponderar hasta qué punto son independientes las confirmaciones.

La posición defendible es estrecha. Las cadenas orales largas transportan demostrablemente señal estructural. Las afirmaciones de que una tradición concreta codifica un acontecimiento datado concreto solo son contrastables allí donde existe una fecha independiente, y su fuerza depende por completo de cuánto restringe la tradición al acontecimiento, que suele ser poco.

En qué punto estamos

Las cadenas largas transportan demostrablemente señal estructural, y eso es todo lo que respalda el resultado filogenético: depende del modelo y no establece nada sobre la antigüedad de un contenido concreto. Que una tradición codifique un acontecimiento datado solo es contrastable donde existe una fecha independiente, y esta literatura es pequeña y su autoría se solapa. Sigue una afirmación más fuerte sobre la escritura.

6. La gran divisoria y la crítica que recibió

Un argumento influyente de mediados del siglo XX sostuvo que la escritura alfabética reestructura el pensamiento y habilita lógica, historia, escepticismo y crítica acumulativa que las culturas orales no pueden sostener (Goody y Watt 1963). Una versión próxima describió oralidad y escritura como modos distintos de conciencia, con la escritura reestructurando la mente que la adquiere (Ong 2013).

La tesis recibió una crítica sostenida desde el trabajo comparativo. La crítica principal sustituyó el modelo autónomo, en el que la escritura es una tecnología uniforme con efectos cognitivos uniformes, por un modelo ideológico, en el que lo que la escritura hace depende de las prácticas sociales en que está inserta, de modo que escrituras distintas en instituciones distintas producen efectos distintos y de la tecnología sola no se sigue ninguna consecuencia general (Street 1984).

Esto importa para la serie y no solo para el tema. La gran divisoria tiene la misma forma que el error que KNO-004 identificó en el estudio de los sistemas de escritura, en otro dominio: una distinción real se convierte en escalera, un lado se trata como destino y todo lo del otro lado se redescubre como carencia. La corrección es la misma en ambos casos. Preguntar qué hace un sistema y para quién, y dejar de preguntar en qué estaba a punto de convertirse.

En qué punto estamos

El argumento de que la escritura alfabética reestructura el pensamiento recibió una crítica comparativa que sustituyó el modelo autónomo por otro donde lo que la escritura hace depende de las prácticas en que está inserta, y de la tecnología sola no se sigue consecuencia general. Eso refuta una afirmación general, no muestra que la escritura no cambie nada. Rechazar el encuadre de la carencia deja en pie costes reales.

7. Coste, y lo que el modo no puede hacer

Rechazar el encuadre de la carencia no significa fingir que no hay límites.

Coste en personas. La fidelidad oral fuerte es cara. Exige especialistas, años de adiestramiento, continuidad institucional y arreglos sociales que premien la transmisión exacta. Cuando esos arreglos fallan, la capacidad desaparece en una generación. Los costes de una y otra son de clases distintas, y este artículo no los compara por unidad de contenido conservado, comparación que exigiría una medida que ninguna de las dos proporciona.

La recuperación suele ser secuencial. El contenido se recupera recorriendo la estructura que lo sostiene, lo cual es eficiente para unidades enteras y malo para la consulta arbitraria. Las tradiciones que incorporan a la estructura recursos de indexación, como enumeraciones fijas o secuencias de lugares, reducen ese coste sin eliminarlo.

Mala para lo arbitrario. La maquinaria de restricción de la sección 3 funciona porque el contenido tiene estructura. Las listas largas de particulares sin motivación, las cantidades exactas y las cadenas arbitrarias son lo que peor estabiliza, y las tradiciones que sí llevan ese material suelen haberle dado antes una estructura artificial. No es casualidad que la escritura más antigua registrara cantidades y nombres.

La pérdida suele ser irreversible. Es la asimetría que distingue con más nitidez la transmisión oral de la gráfica. Un manuscrito perdido puede recuperarse, estar copiado en otro sitio o excavar. Una tradición que nunca se registró no deja nada cuando cesan sus últimos hablantes; donde hubo grabación o transcripción, lo que sobrevive es un registro de la tradición y no la tradición. La amenaza a las lenguas no es por tanto solo una cuestión lingüística: el análisis de redes de lenguas amenazadas identifica concentraciones geográficas donde la pérdida está estructuralmente impulsada por patrones de contacto y no repartida de forma uniforme (Lee, Siew y Ng 2022), y cada pérdida se lleva consigo cuanto se conservara solo en esa lengua y solo en la memoria.

En qué punto estamos

La fidelidad oral fuerte es cara en especialistas, adiestramiento y continuidad institucional, y la capacidad desaparece en una generación cuando esos arreglos fallan. La recuperación es secuencial, los particulares arbitrarios son lo que peor estabiliza y la pérdida suele ser irreversible. El artículo no compara los dos modos por unidad de contenido conservado, porque ninguno da esa medida. Parte de esto la evidencia no lo resuelve en absoluto.

8. Lo que la evidencia no puede resolver

Fidelidad anterior a la observación. La fidelidad medida en tradiciones orales procede de tradiciones observadas por investigadores, es decir de tradiciones que sobrevivieron para ser observadas. Nada corrige por las prácticas que fracasaron, de modo que las estimaciones de lo que la transmisión oral puede lograr están sesgadas al alza por supervivencia.

Antigüedad de un contenido concreto. Los métodos estructurales datan linajes, no enunciados. Un tipo de relato puede ser antiguo mientras cada palabra de una versión registrada es reciente, y ningún método distingue hoy un elemento antiguo de uno nuevo bien integrado dentro de una tradición viva.

La cuestión de la literalidad en tradiciones no documentadas. Que una tradición transmitiera literalmente solo puede establecerse allí donde se registraron varias versiones independientes antes de que la tradición se alterara. Para la mayor parte del registro histórico esa condición no se cumple.

Las categorías. Como en KNO-004, los seis modos de la sección 2 son instrumentos analíticos. Las prácticas reales los mezclan, y una misma comunidad puede mover contenido entre modos de forma deliberada.

En qué punto estamos

Quedan cuatro límites. La fidelidad medida procede solo de tradiciones que sobrevivieron para ser observadas, así que las estimaciones están sesgadas al alza. Los métodos estructurales datan linajes y no enunciados. La transmisión literal solo es demostrable donde se registraron versiones independientes antes de perturbar la tradición. Los seis modos son instrumentos que las prácticas reales mezclan. Lo que sobrevive a esos límites es lo que el artículo concluye.

9. Conclusión

La transmisión oral no es lo que precedió a la escritura. Es una familia de técnicas para sostener conocimiento en una población, que va desde prácticas casi sin control de errores hasta prácticas con un control de errores que la escritura no mejoró durante mucho tiempo. Sus formas fuertes logran su fidelidad restringiendo lo que puede decirse, y por eso favorecen verso, fórmula y canto, y por eso fallan con el detalle arbitrario.

Tratarla como etapa produce dos errores a la vez. Halaga a la escritura, al acreditarle la solución de un problema que los regímenes orales fuertes ya habían resuelto para su propio contenido, y oscurece lo que la escritura cambió de verdad, que fue coste, portabilidad, acceso aleatorio y capacidad de registrar lo arbitrario. Son cambios grandes. No son lo mismo que hacer posible el pensamiento.

En qué punto estamos

La transmisión oral es una familia de técnicas y no una etapa previa a la escritura, y sus formas fuertes compran fidelidad restringiendo lo que puede decirse. Tratarla como etapa halaga a la escritura y oscurece lo que la escritura sí cambió: coste, portabilidad, acceso aleatorio y lo arbitrario. Son cambios grandes que no equivalen a hacer posible el pensamiento. Qué sabemos, y cómo lo sabemos, es lo que sigue.

Lo que sabemos y cómo lo sabemos

El trabajo cognitivo sobre las tradiciones fuertes establece que la estabilidad procede de muchas restricciones individualmente débiles actuando a la vez — metro, melodía, fórmula, imagen y secuencia narrativa — que juntas dejan a menudo una sola opción donde cada una por separado deja muchas. El trabajo de campo de Parry y Lord muestra a un cantor que produce la misma trama con extensión parecida y palabras muy distintas sin vivirlo como un fallo. La recitación permutada védica muestra a una comunidad haciendo el error detectable y no solo improbable.

Lo que todavía no sabemos

La fidelidad medida procede solo de tradiciones que sobrevivieron para ser observadas, y nada corrige por las prácticas que fallaron, así que las estimaciones de lo que puede lograr la transmisión oral están sesgadas al alza. Los métodos estructurales datan linajes y no enunciados, de modo que un tipo de relato puede ser antiguo mientras cada palabra registrada es reciente. Si una tradición transmitió literalmente solo puede resolverse donde se registraron versiones independientes antes de perturbarla.

Qué podría hacernos cambiar de opinión

Versiones independientes de una tradición registradas antes de perturbarla llevarían la cuestión de la literalidad del método a la evidencia. Un método capaz de distinguir un elemento antiguo de uno nuevo bien integrado dentro de una tradición viva haría lo mismo con la antigüedad del contenido. Los resultados filogenéticos dependen del modelo, y otras codificaciones discutidas en esa literatura podrían revisar hasta dónde llega la señal.

Si solo recuerdas una cosa

La fidelidad oral fuerte se produce restringiendo lo que puede decirse y no esforzándose más por recordarlo, y por eso favorece el verso y falla en el detalle arbitrario.

Artículos relacionados

- KNO-003 · *De las marcas a los símbolos*
- KNO-004 · *De los símbolos a la escritura*

Referencias

Franks, Leigh, Patrick D. Nunn, y Adrian McCallum. 2024. «The selective geography of volcanism in oral traditions». *Quaternary Research* 122: 1–17. DOI 10.1017/qua.2024.29. Franks, Leigh, Patrick D. Nunn, Adrian McCallum, y Gertz, Janine. 2025. «Stories in Stone». *Geoheritage* 17 (2): art. 73. DOI 10.1007/s12371-025-01116-2. Goody, Jack, y Ian Watt. 1963. «The Consequences of Literacy». *Comparative Studies in Society and History* 5 (3): 304–345. DOI 10.1017/S0010417500001730. Graça da Silva, Sara, y Jamshid J. Tehrani. 2016. «Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales». *Royal Society Open Science* 3 (1): art. 150645. DOI 10.1098/rsos.150645. Henige, David. 2009. «Impossible to Disprove Yet Impossible to Believe: The Unforgiving Epistemology of Deep-Time Oral Tradition». *History in Africa* 36: 127–234. DOI 10.1353/hia.2010.0014. Lee, Nala H., Cynthia S. Q. Siew, y Nadine H. N. Ng. 2022. «The network nature of language endangerment hotspots». *Scientific Reports* 12: art. 10803. DOI 10.1038/s41598-022-14479-1. Lord, Albert B. 2019. *The Singer of Tales*. Ed. David F. Elmer. 3.^a ed. Cambridge, MA: Harvard University Press. ISBN 978-0-674-97573-6. Primera edición 1960. Nunn, Patrick D., y Nicholas J. Reid. 2016. «Aboriginal Memories of Inundation of the Australian Coast Dating from More than 7000 Years Ago». *Australian Geographer* 47 (1). DOI 10.1080/00049182.2015.1077539. Ong, Walter J. 2013. *Orality and Literacy: The Technologizing of the Word*. 30.^o aniversario, con dos ensayos de John Hartley. London: Routledge. ISBN 978-0-415-53838-1. Rubin, David C. 1995. *Memory in Oral Traditions: The Cognitive Psychology of Epic, Ballads, and Counting-out Rhymes*. New York: Oxford University Press. ISBN 978-0-19-508211-1. Rústica 1997, ISBN 978-0-19-512032-5. Staal, Frits. 1986. *The Fidelity of Oral Tradition and the Origins of Science*. Amsterdam: North-Holland / Koninklijke Nederlandse Akademie van Wetenschappen. ISBN 0-444-85669-2. Street, Brian V. 1984. *Literacy in Theory and Practice*. Cambridge Studies in Oral and Literate Culture 9. Cambridge: Cambridge University Press. ISBN 0-521-24785-3.

KNO-006 · Mapas y representación espacial

KNO · Formas de conocer

Juan Segurola

v0.7 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia en publicación:** Todos los derechos reservados **Título bilingüe:** Maps and Spatial Representation / Mapas y representación espacial

Resumen del artículo

¿Qué es un mapa y cuánto del conocimiento espacial humano cubre realmente esa palabra? Este artículo separa tres cosas que suelen fundirse en una: la capacidad de orientarse y navegar, el artefacto externo que representa relaciones espaciales y las prácticas llamadas cartografía. Pregunta qué establecen y qué no los hallazgos sobre células de lugar y de retícula respecto del formato del pensamiento espacial. Examina las cartas de oleaje y los marcos de referencia móviles, los candidatos prehistóricos en disputa y la restricción geométrica que obliga a todo mapa plano de una esfera a distorsionar algo. Después pregunta qué se sigue para la lectura de los mapas como argumentos una vez que la distorsión y la selección son inevitables, por qué el dominio de una proyección en línea no disuelve la restricción, y cuáles de las preguntas planteadas no puede resolver la evidencia reunida aquí.

En 60 segundos

El conocimiento espacial, la representación espacial externa y la cartografía son tres cosas separables, y un navegante experto puede poseer la primera en grado alto sin usar ningún artefacto gráfico durante la travesía. Las células de lugar y de retícula son hallazgos robustos sobre actividad neuronal, pero no establecen que el pensamiento espacial humano adopte la forma de un mapa gráfico; el conocimiento de ruta puede sostener viajes exitosos sin ser un mapa de conjunto degradado. Las cartas de oleaje marshalesas y la técnica de las Carolinas conocida como etak muestran que el razonamiento espacial puede organizarse en torno a modelos del comportamiento del océano y a marcos de referencia distintos del marco fijo de un plano a escala. La geometría impone después un límite duro: como la esfera y el plano difieren en curvatura gaussiana, ningún mapa plano de una esfera puede conservarlo todo, de modo que una proyección codifica una decisión sobre qué distorsiones aceptar. Aparte de eso, todo mapa selecciona, de modo que la selección, la escala, la orientación y la clasificación son decisiones editoriales y no registro neutral, y Web Mercator muestra la restricción geométrica persistiendo en los medios interactivos.

Por qué importa esta pregunta

Un mapa puede leerse como una descripción transparente del espacio. Un atlas de carreteras, una carta meteorológica o un mapamundi en la pared de un aula se leen como imagen de lo que hay, y esa lectura suele ser inofensiva hasta que el mapa sirve para comparar tamaños, inferir distancias o zanjar una discusión sobre el mundo. Saber que todo mapa plano del globo distorsiona algo, y que el tipo de distorsión que lleva procede de una elección y no de un descuido, cambia lo que un lector puede concluir con responsabilidad a partir de él. Separar los mapas del conocimiento espacial cambia también qué sistemas espaciales cuentan como conocimiento.

1. Tres cosas que no son lo mismo

El conocimiento espacial es la capacidad de orientarse, localizar o navegar. Una representación espacial externa es un artefacto que está en lugar de relaciones espaciales. La cartografía es el estudio y la práctica de hacer y usar mapas bajo convenciones explícitas o implícitas, con historias en muchas sociedades. Los tres están relacionados, pero no son idénticos.

Un navegante experto que cruza mar abierto sin instrumentos posee conocimiento espacial de alto orden y puede no usar ningún artefacto gráfico durante la travesía. Una carta de instrucción marshalesa puede destacar el comportamiento del oleaje en vez de funcionar como un plano a escala. Un atlas de carreteras moderno es una representación externa cuyo usuario puede conocer poco del terreno que cubre. Nada obliga a las tres capacidades a coincidir.

El artículo usa *mapa* en sentido amplio para un artefacto externo que representa relaciones espaciales. *Mapa proyectado* designa, de forma más restringida, una representación que transforma ubicaciones de una superficie curva a un plano. Ninguno de los términos se reserva a una tradición europea.

En qué punto estamos

Las tres pueden separarse, como muestran los ejemplos, de modo que nada autoriza a tratarlas como una sola cosa con tres nombres. El artículo fija su vocabulario en lugar de descubrirlo: mapa cubre cualquier artefacto externo que represente relaciones espaciales. Si el pensamiento espacial humano se parece a un mapa gráfico es otra cuestión empírica, y la sección siguiente pregunta qué sostiene realmente la neurociencia.

2. Qué hace el cerebro con el espacio

Dos hallazgos anclan el lado biológico y ninguno de los dos dice lo que suele entenderse por la expresión popular mapa cognitivo.

Las células de lugar del hipocampo se activan cuando un animal ocupa una posición determinada. O'Keefe y Nadel propusieron que el hipocampo sostiene una representación aloécéntrica organizada con respecto al entorno (O'Keefe y Nadel 1978). Aquí, aloécéntrica resume el marco no egocéntrico de la obra; no se presenta como una cita literal de pp. 89–90. Más tarde se observó que, en ratas, ciertas células de la corteza entorrinal medial dorsocaudal se activaban en los vértices de una retícula triangular periódica a lo largo del entorno experimental (Hafting et al. 2005). Su función en un sistema espacial más general es una interpretación respaldada por ese patrón de activación, pero no idéntica a él.

Lo que estos hallazgos no establecen es que el cerebro construya un mapa en el sentido gráfico cotidiano. Un modelo clásico del desarrollo distingue conocimiento de hitos, conocimiento de ruta organizado en secuencias y conocimiento de conjunto que integra relaciones espaciales más allá de una ruta (Siegel y White 1975). La secuencia propuesta ha sido influyente, pero no debe convertirse en una ley universal del desarrollo. El conocimiento de ruta puede permitir desplazamientos eficaces sin ser un mapa de conjunto degradado, y su flexibilidad varía según el aprendizaje y la tarea.

La relevancia para el resto del artículo es directa. Si la competencia espacial humana descansa en buena medida en conocimiento secuenciado y dependiente del punto de vista, entonces la representación de conjunto que los europeos acabaron llamando mapa no es la externalización natural de cómo la gente piensa el espacio. Es una opción entre varias, y las otras están bien atestiguadas.

En qué punto estamos

Las células de lugar y de retícula son hallazgos robustos sobre actividad neuronal y no muestran que el pensamiento espacial adopte forma gráfica. La distinción entre hitos, ruta y conjunto sigue siendo influyente sin ser ley universal del desarrollo. Si el conocimiento secuenciado y dependiente del punto de vista carga buena parte de la competencia humana, el mapa de conjunto es una externalización entre varias y no la natural.

3. Representaciones que no son imágenes del terreno

3.1 Cartas de oleaje

Las cartas de varillas marshalesas son artefactos de instrucción y enseñanza asociados con la navegación por el oleaje. Sus elementos pueden codificar relaciones entre islas y patrones de oleaje, incluidas las formas en que las islas alteran las olas; los distintos tipos de carta no cumplen todos la misma función. Por ello no equivalen a planos geográficos a escala, pero sería demasiado amplio afirmar que nunca representan la disposición de las islas (Genz et al. 2009; Genz 2016). En el mar, los navegantes detectan patrones relevantes en parte mediante el movimiento de la canoa.

Lo importante no es que las cartas carezcan de contenido espacial, sino que ese contenido es inseparable de un modelo del comportamiento oceánico y de la práctica de navegación. Leerlas como mapas topográficos ordinarios hace perder la función para la que se enseñaban.

3.2 Marcos de referencia que se mueven

La navegación de las islas Carolinas incluye una técnica conocida como *etak*, en la que el avance puede concebirse mediante islas de referencia que pasan por sucesivas demoras estelares. El sistema se describe en la literatura etnográfica sobre Puluwat y en trabajos comparados sobre navegación del Pacífico (Gladwin 1995; Lewis 1994).

El punto analítico es que el razonamiento espacial puede organizar el movimiento alrededor de distintos marcos de referencia. La descripción etnográfica establece la práctica; no establece por sí sola que un marco sea computacionalmente óptimo ni que toda una tradición europea emplee únicamente su opuesto.

3.3 El punto general

Los tres libros del volumen 2 de *The History of Cartography* examinan las tradiciones islámica y del sur de Asia, de Asia oriental y sudoriental, africana, americana, ártica, australiana y del Pacífico en sus propios contextos históricos (Harley y Woodward 1992; 1994; Woodward y Lewis 1998). Su organización permite comparar sin tratar una tradición como plantilla universal.

En qué punto estamos

Un artefacto puede portar contenido espacial y rechazar la forma de un plano a escala, y leerlo como mapa topográfico descarta la función para la que fue hecho. Una carta de oleaje es inseparable de un modelo del comportamiento del océano, y el etak puede organizar el movimiento en torno a un marco que se desplaza. Los artefactos candidatos más antiguos plantean un problema de identificación más difícil.

4. Los mapas candidatos más antiguos están en disputa

La pintura mural de Çatalhöyük suele presentarse como el mapa más antiguo conocido, con un plano de la aldea bajo un volcán en erupción. La lectura ha sido discutida en detalle, con el argumento de que la composición se explica mejor como una piel de leopardo y un diseño geométrico que como un plano y una montaña (Meece 2006). Un estudio posterior aceptó una lectura representacional del elemento superior y sostuvo, por razones geocronológicas, que puede representar una erupción, sin zanjar si el elemento inferior es un plano (Schmitt et al. 2014).

El riesgo metodológico es la circularidad: la semejanza con una categoría moderna motiva una identificación, y después la identificación se cita como prueba de que la categoría es antigua. La evidencia contextual independiente puede fortalecer la inferencia, pero la semejanza por sí sola no puede establecer la función.

Esto no es escepticismo sobre la representación espacial antigua, para la que sobra evidencia posterior inequívoca. Es escepticismo sobre el movimiento concreto que convierte una imagen sugerente en el primer mapa.

En qué punto estamos

Identificar una imagen prehistórica como mapa solo por semejanza invita a la circularidad: la categoría moderna motiva la lectura, y la lectura se ofrece luego como prueba de que la categoría es antigua. La evidencia contextual independiente puede reforzarla; la semejanza sola no fija la función. La duda es estrecha y no alcanza a la representación espacial antigua en general. Lo que sigue no depende de la historia.

5. Por qué ningún mapa plano de una esfera puede ser fiel

Aquí el argumento deja de ser histórico.

Una esfera y un plano tienen distinta curvatura gaussiana, y las isometrías locales conservan la curvatura gaussiana, el Theorema Egregium (do Carmo 1976, p. 234). De ello deduce este artículo que ninguna región abierta de una esfera puede proyectarse isométricamente sobre un plano: alguna propiedad métrica debe distorsionarse. La fuente plantea expresamente la imposibilidad esfera-plano como ejercicio 4 en p. 237; el desarrollo demostrativo expuesto aquí corresponde al artículo.

Ese resultado convierte lo que parece un problema de ingeniería en una restricción. Todo mapa plano de una superficie con curvatura gaussiana no nula, incluida la esfera, distorsiona. Una proyección puede conservar ángulos, o áreas, o distancias a lo largo de ciertas líneas, o un compromiso entre ellas, y no puede conservarlo todo. El cartógrafo debe, por tanto, elegir qué distorsiones acepta, y la exactitud sigue siendo evaluable según el propósito previsto del mapa.

La proyección de Mercator es la ilustración habitual. En su modelo esférico es conforme y una línea de rumbo constante es recta; su escala y la inflación del área aumentan hacia los polos. Monmonier documenta tanto su valor para la navegación como su uso posterior en mapas murales y otros contextos no navegacionales, donde la compensación puede ajustarse mal al propósito (Monmonier 2004).

La lección general es que una proyección codifica una decisión sobre qué importa, y que quien no conoce la decisión leerá mal el mapa en una dirección que el mapa no declara.

En qué punto estamos

Como esfera y plano difieren en curvatura gaussiana, ninguna región abierta de una se tiende isométricamente sobre el otro, así que todo mapa plano de una esfera distorsiona algo. Esa consecuencia la deduce este artículo del Theorema Egregium, no la cita del libro. Lo que parece un límite de ingeniería es una restricción: las proyecciones distribuyen de modo distinto la distorsión inevitable, y la exactitud debe evaluarse según el propósito.

6. Los mapas como argumentos

Una vez que la distorsión es inevitable y la selección necesaria, la afirmación de que un mapa se limita a mostrar lo que hay se vuelve insostenible. La literatura de cartografía crítica hace de esto su punto central: los mapas son artefactos retóricos cuya aparente neutralidad es ella misma un efecto que producen (Harley 1989), y se ha argumentado que la forma cartográfica puede naturalizar los intereses a los que un mapa sirve en vez de enunciarlos (Wood 1992; Wood, Fels y Krygier 2010).

Cuatro decisiones editoriales recurrentes ayudan a exponer esa selectividad.

Selección. Qué está en el mapa. Todo mapa omite casi todo, y las omisiones no son aleatorias.

Escala. Qué es demasiado pequeño para aparecer. La escala es una afirmación sobre qué cuenta como rasgo.

Orientación. El norte arriba es una convención con historia, no una propiedad de la Tierra. Los mapas pueden adoptar otras orientaciones según su finalidad y tradición.

Clasificación. Qué cuenta como la misma clase de cosa, lo que importa una taxonomía que el mapa muestra después como si fuera terreno.

Esto no hace los mapas arbitrarios ni poco fiables, y la literatura crítica se lee a menudo mal como si lo dijera. Un mapa puede comprobarse, corregirse y usarse, y la geometría de la sección 5 no es cuestión de opinión. Lo que se sigue es más estrecho: la neutralidad que un mapa proyecta es una propiedad del género, y evaluar un mapa exige conocer su fin.

En qué punto estamos

Si la distorsión es inevitable y la selección necesaria, un mapa no puede limitarse a mostrar lo que hay. Selección, escala, orientación y clasificación son decisiones recurrentes por las que opera esa selectividad, y la neutralidad que un mapa proyecta es un efecto del género, no una propiedad del terreno. Esto no vuelve arbitrarios los mapas, y la geometría se sostiene al margen de la crítica.

7. El caso digital

La cartografía interactiva en línea ha convertido Web Mercator en la proyección de visualización dominante. Battersby y sus colegas documentan ese predominio y analizan sus ventajas técnicas, limitaciones cartográficas e implicaciones cognitivas; no establecen una única causa exclusiva para su adopción (Battersby et al. 2014). Web Mercator usa las ecuaciones del Mercator esférico, se adapta bien a la visualización por teselas y conserva la intensa inflación de área en latitudes altas propia de las vistas mundiales de tipo Mercator.

El caso muestra que la restricción geométrica de la sección 5 no desaparece cuando el medio se vuelve interactivo. La arquitectura técnica, la interoperabilidad, los estándares heredados y las decisiones de diseño pueden influir en la proyección que encuentran los usuarios. Asignar su peso causal relativo exige evidencia histórica adicional a las propiedades matemáticas de la proyección.

En qué punto estamos

El dominio de Web Mercator en línea muestra la restricción geométrica sobreviviendo al paso a los medios interactivos, con la misma inflación severa de áreas en latitudes altas. Arquitectura técnica, interoperabilidad, estándares heredados y diseño pueden moldear qué proyección encuentra el usuario, y el análisis citado no aísla causa exclusiva. Asignar pesos causales exigiría evidencia histórica que las matemáticas no dan.

8. Lo que la evidencia no puede resolver

Si un artefacto dado es un mapa. Sin evidencia independiente de uso, identificar una imagen prehistórica como mapa es una inferencia por semejanza, y la semejanza con una categoría moderna es evidencia débil sobre una antigua.

Qué son los mapas cognitivos. Las células de lugar y las de retícula son hallazgos robustos sobre actividad neuronal. El paso de ahí a una afirmación sobre el formato del pensamiento espacial humano es una interpretación, y la distinción entre ruta y conjunto sugiere que el formato varía según la tarea y la persona.

Si los sistemas no cartográficos son mapas. Que una carta de oleaje o una ruta cantada sea un mapa depende de la definición adoptada, y este artículo no pretende que su propia estipulación resuelva una cuestión en parte terminológica.

La intención tras la selección. Que los mapas pueden codificar intereses se argumenta en la literatura crítica. Que un mapa concreto codifique una intención concreta exige evidencia sobre quienes lo hicieron, y la literatura crítica es más fuerte en la afirmación general que en las atribuciones individuales.

En qué punto estamos

Cuatro preguntas resisten la evidencia reunida: si un artefacto dado se usó como mapa, qué formato adopta el pensamiento espacial, si los sistemas no cartográficos caen bajo el término y qué intención codifica un mapa concreto. Una es en parte terminológica más que empírica, y el vocabulario estipulado aquí no la resuelve. La síntesis final expone lo que sí se sigue.

9. Conclusión

El espacio puede conocerse sin artefacto gráfico, representarse sin un plano del terreno a escala y proyectarse de una esfera sobre un plano solo aceptando compensaciones geométricas. Un mapa cartográfico es una familia potente de soluciones, no la forma universal del conocimiento espacial.

El error recurrente es tratar una solución como el caso general y todo lo demás como aproximaciones a ella. Un modelo de oleaje, una secuencia estelar, una descripción de ruta y una retícula proyectada responden a preguntas distintas sobre el espacio. Un mapa mundial conforme acepta distorsión de área para conservar los ángulos locales; esa compensación es lo bastante precisa sin llamar mentira al mapa.

En qué punto estamos

Conocer el espacio no exige artefacto gráfico y representarlo no exige imagen a escala, pero una esfera solo se proyecta sobre un plano con compromisos geométricos. El mapa cartográfico es por tanto una familia potente de soluciones, no la forma general del conocimiento espacial. El error recurrente toma esa familia por el patrón y lee todo lo demás como aproximación, cuando las alternativas responden a preguntas distintas.

Nota de frontera

Con KNO-005, Lenguaje oral y transmisión. El conocimiento espacial sostenido y transmitido oralmente, incluidas la instrucción de ruta y las rutas cantadas, se trata en KNO-005 como instancia de transmisión restringida. Este artículo retoma la cuestión representacional que KNO-005 dejó expresamente aplazada: qué hace un artefacto externo cuando está en lugar de relaciones espaciales. Los mecanismos de transmisión, la fidelidad y la memoria no se reexaminan aquí.

Con KNO-007, Calendarios y medida del tiempo. La representación del tiempo, la división de ciclos y los instrumentos de medida pertenecen a KNO-007. Donde este artículo roza el tiempo, en la navegación por secuencia estelar y en la estima, lo hace solo como elemento de método espacial. La cuestión general de cómo un proceso periódico se convierte en medida es de KNO-007.

Con KNO-004 y KNO-013. La notación y la codificación gráfica de la lengua pertenecen a KNO-004. La cuestión más amplia de qué es un modelo y cómo se relacionan los modelos con la realidad es de KNO-013, y este artículo no la anticipa: trata la proyección como restricción geométrica y no como caso de estudio sobre modelización.

Lo que sabemos y cómo lo sabemos

- Las células de retícula son un hallazgo establecido del registro unitario en ratas (Hafting et al. 2005). O'Keefe y Nadel propusieron que el hipocampo sostiene una representación alocéntrica (O'Keefe y Nadel 1978).
- Las isometrías locales conservan la curvatura gaussiana, el Theorema Egregium (do Carmo 1976, p. 234). De él deduce este artículo que ninguna región abierta de una esfera se proyecta isométricamente sobre un plano.
- El valor navegacional de Mercator y su uso posterior no navegacional están ambos documentados (Monmonier 2004).
- El dominio de Web Mercator en la cartografía en línea, sus ventajas técnicas y sus limitaciones están documentados, sin que se establezca una causa exclusiva de adopción (Battersby et al. 2014).
- Los volúmenes comparativos cubren tradiciones islámicas, del sur de Asia, de Asia oriental y sudoriental, africanas, americanas, árticas, australianas y del Pacífico en sus propios contextos (Harley y Woodward 1992; 1994; Woodward y Lewis 1998).

Lo que todavía no sabemos

Si un artefacto prehistórico dado se hizo o se usó como mapa sigue abierto, porque la semejanza con una categoría moderna es evidencia débil sobre una antigua. Qué formato adopta el pensamiento espacial humano no lo resuelven los hallazgos neuronales, y la distinción entre ruta y conjunto sugiere que varía según tarea y persona. Si una carta de oleaje o una ruta cantada cuenta como mapa es en parte terminológico. Y atribuir una intención concreta a un mapa concreto exige evidencia sobre sus autores que el argumento general no proporciona.

Qué podría hacernos cambiar de opinión

Evidencia contextual independiente de uso, y no de semejanza, podría resolver varias identificaciones en disputa en un sentido o en otro. Un trabajo conductual o neuronal que mostrara que la representación de conjunto es el formato por defecto del pensamiento espacial humano, y no uno entre varios, debilitaría la tesis del artículo de que el mapa de conjunto no es la externalización natural. Evidencia histórica que aislara por qué se adoptó Web Mercator convertiría una lista de factores plausibles en una explicación causal. La restricción geométrica de la sección 5 no admite esta clase de revisión.

Si solo recuerdas una cosa

Ningún mapa plano de una esfera conserva todas las propiedades métricas, así que la pregunta que cabe hacerle no es si distorsiona, sino qué distorsiones acepta.

Artículos relacionados

- KNO-004 · De los símbolos a la escritura
- KNO-005 · Lenguaje oral y transmisión
- KNO-007 · Calendarios y medida del tiempo

Referencias

- Battersby, Sarah E., Michael P. Finn, E. Lynn Usery, y Kristina H. Yamamoto. 2014. «Implications of Web Mercator and Its Use in Online Mapping». *Cartographica* 49 (2): 85–101. DOI 10.3138/carto.49.2.2313.
- do Carmo, Manfredo P. 1976. *Differential Geometry of Curves and Surfaces*. Englewood Cliffs: Prentice-Hall. ISBN 978-0-13-212589-5.
- Genz, Joseph H. 2016. «Resolving Ambivalence in Marshallese Navigation: Relearning, Reinterpreting, and Reviving the “Stick Chart” Wave Models». *Structure and Dynamics* 9 (1). DOI 10.5070/SD991031888.
- Genz, Joseph, Jerome Aucan, Mark Merrifield, Ben Finney, Korent Joel, y Alson Kelen. 2009. «Wave Navigation in the Marshall Islands: Comparing Indigenous and Western Scientific Knowledge of the Ocean». *Oceanography* 22 (2): 234–245. DOI 10.5670/oceanog.2009.52.
- Gladwin, Thomas. 1995. *East Is a Big Bird: Navigation and Logic on Puluwat Atoll*. Cambridge, MA: Harvard University Press. ISBN 978-0-674-22426-1.
- Hafting, Torkel, Marianne Fyhn, Sturla Molden, May-Britt Moser, y Edvard I. Moser. 2005. «Microstructure of a Spatial Map in the Entorhinal Cortex». *Nature* 436 (7052): 801–806. DOI 10.1038/nature03721.
- Harley, J. B. 1989. «Deconstructing the Map». *Cartographica* 26 (2): 1–20. DOI 10.3138/E635-7827-1757-9T53.
- Harley, J. B., y David Woodward, eds. 1992. *The History of Cartography, Volume 2, Book 1: Cartography in the Traditional Islamic and South Asian Societies*. Chicago: University of Chicago Press. ISBN 0-226-31635-1.
- Harley, J. B., y David Woodward, eds. 1994. *The History of Cartography, Volume 2, Book 2: Cartography in the Traditional East and Southeast Asian Societies*. Chicago: University of Chicago Press. ISBN 0-226-31637-8.
- Lewis, David. 1994. *We, the Navigators: The Ancient Art of Landfinding in the Pacific*. 2.^a ed. Editado por Derek Oulton. Honolulu: University of Hawai'i Press. ISBN 978-0-8248-1582-0.
- Meece, Stephanie. 2006. «A Bird's Eye View – of a Leopard's Spots: The Çatalhöyük “Map” and the Development of Cartographic Representation in Prehistory». *Anatolian Studies* 56: 1–16. DOI 10.1017/S0066154600000727.
- Monmonier, Mark. 2004. *Rhumb Lines and Map Wars: A Social History of the Mercator Projection*. Chicago: University of Chicago Press. ISBN 978-0-226-53431-2.
- O'Keefe, John, y Lynn Nadel. 1978. *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press. ISBN 978-0-19-857206-0.
- Schmitt, Axel K., Martin Daniščík, Erkan Aydar, Erdal Şen, İnan Ulusoy, y Oscar M. Lovera. 2014. «Identifying the Volcanic Eruption Depicted in a Neolithic Painting at Çatalhöyük, Central Anatolia, Turkey». *PLOS ONE* 9 (1): e84711. DOI 10.1371/journal.pone.0084711.

Siegel, Alexander W., y S. H. White. 1975. «The Development of Spatial Representations of Large-Scale Environments». *Advances in Child Development and Behavior* 10: 9–55. DOI 10.1016/S0065-2407(08)60007-5.

Wood, Denis. 1992. *The Power of Maps*. New York: Guilford Press. ISBN 978-0-89862-493-9.

Wood, Denis, con John Fels y John Krygier. 2010. *Rethinking the Power of Maps*. New York: Guilford Press. ISBN 978-1-59385-366-2.

Woodward, David, y G. Malcolm Lewis, eds. 1998. *The History of Cartography, Volume 2, Book 3: Cartography in the Traditional African, American, Arctic, Australian, and Pacific Societies*. Chicago: University of Chicago Press. ISBN 0-226-90728-7.

KNO-007 · Calendarios y medida del tiempo

KNO · Formas de conocer

Juan Segurola

v0.2 · ES

Todos los derechos reservados

Resumen del artículo

Casi todo cómputo del tiempo se apoya en tres ciclos observables que no se dividen entre sí. Para la época 2000 el año trópico medio es de 365,24219 días y el mes sinódico medio de 29,53059 días, ambos medias aproximadas que cambian lentamente, de modo que un calendario que busque alineación lunar o estacional necesita reglas, observación, intercalación o deriva aceptada. KNO-007 pregunta qué se sigue. Compara estrategias lunares, solares y lunisolares sin ordenarlas como etapas de refinamiento, realiza su propia aritmética y la marca como propia, y muestra que un ciclo calendárico no tiene por qué seguir un periodo astronómico. Separa después el segundo del SI, definido por una transición del cesio desde 1967, de las escalas de tiempo civil que la convención mantiene cerca de una Tierra que gira. Cierra con candidatos prehistóricos discutidos, con qué coordina un calendario y con los límites de la evidencia.

En 60 segundos

Doce meses sinódicos medios suman 354,36708 días, 10,87511 menos que un año trópico, así que un calendario lunar recorre las estaciones y el islámico lo hace por diseño y no por fracaso. Suelta el mes lunar y conserva el año y aparece el año civil egipcio: doce meses de treinta días más cinco epagómenos, 365 en total, sin día intercalado y con una deriva propia más lenta. Consérvalos ambos y el precio es la intercalación, porque diecinueve años trópicos y 235 meses sinódicos difieren en 0,08704 días, unas 2,09 horas, bastante poco para levantar sobre ello un ciclo de diecinueve años y no lo bastante para hacerlo exacto. El año juliano se pasa en 0,00781 días y el gregoriano en 0,00031, así que 1582 redujo un residuo en vez de eliminarlo, y revisó además las tablas lunares de la Pascua. Entretanto el segundo dejó de ser astronómico en 1967 y el tiempo civil no.

Por qué importa esta pregunta

Leída como una escalera, esta historia se convierte en una marcha hacia el año gregoriano con todo lo demás como intento fallido, y la pregunta real desaparece. Un calendario lunar no es un calendario solar inexacto, y una cuenta de 260 días no es un año inexacto: los ciclos de trece y veinte días no tienen base astronómica directa sobre la que ser inexactos. Un calendario es además un dispositivo de coordinación, y por eso reformar uno nunca es solo un cálculo, y por eso la decisión de 2022 de dejar que el tiempo atómico y el rotacional se separen más se argumentó por el riesgo para la infraestructura digital crítica.

1. Tres periodos que no encajan

Muchos sistemas de cómputo del tiempo se apoyan en tres ciclos observables: el día y la noche, las fases lunares y el retorno estacional.

Para la época 2000, una obra astronómica de referencia da un año trópico medio de 365,24219 días y un mes sinódico medio de 29,53059 días, tomando cada día como 86.400 segundos SI (Richards 2013, en Urban y Seidelmann, eds.). La misma fuente subraya que son periodos medios aproximados y que cambian lentamente con el tiempo.

De esas dos cifras se siguen tres consecuencias por pura aritmética. **El artículo realiza esa aritmética por sí mismo; los productos siguientes son deducciones de los valores citados y no citas.**

- Doce meses sinódicos suman 354,36708 días, **10,87511 días menos que un año trópico.**
- El mes sinódico medio no es un número entero de días de 86.400 segundos, por lo que meses calendáricos de longitud fija no pueden permanecer exactamente alineados con todas las lunaciones.
- El año trópico medio no es un número entero de meses sinódicos medios, por lo que una secuencia fija de meses lunares no puede permanecer exactamente alineada con las estaciones sin ajustes.

Esto no es defecto de un calendario concreto. Es consecuencia de las relaciones aproximadas y cambiantes entre periodos astronómicos. Los calendarios que buscan alineación lunar o estacional deben usar reglas, observaciones, intercalaciones o aceptar deriva. Un calendario no tiene por qué buscar ninguna de esas alineaciones, como muestra la sección 4.

Existe un paralelo útil pero limitado con KNO-006. Un mapa proyectado afronta un teorema geométrico exacto; un calendario afronta periodos medidos empíricamente que varían con el tiempo. Ambos hacen visibles compensaciones, pero el estatuto matemático de las dos restricciones no es el mismo.

En qué punto estamos

Tres ciclos observables sostienen el cómputo del tiempo, y los dos periodos medidos son periodos medios aproximados que cambian lentamente. De ahí se sigue que meses de longitud fija no pueden permanecer alineados con todas las lunaciones, ni una secuencia fija de ellos con las estaciones sin ajuste. Esa aritmética se hace aquí y no se cita, y qué hace un calendario con una restricción empírica es una elección.

2. Tres estrategias extendidas

2.1 Conservar el mes, perder la estación

Un calendario de doce meses sinódicos medios queda unos once días por debajo de un año trópico, así que sus fechas se desplazan por las estaciones. Con las cifras anteriores, el desplazamiento acumulado equivale aproximadamente a un año trópico tras 33,6 años lunares.

El calendario islámico es lunar: mantiene los meses en relación con las lunaciones y no mantiene su año sincronizado con el Sol (Richards 2013). Su deriva estacional es, por tanto, consecuencia de la estructura elegida y no prueba de que el calendario haya fracasado como calendario solar.

2.2 Conservar la estación, perder la luna

La elección opuesta abandona el mes lunar como unidad astronómica y conserva solo su nombre. El año civil egipcio es el caso temprano más claro: doce meses de treinta días más cinco días epagómenos, trescientos sesenta y cinco en total, descritos en el estudio de referencia como un año sin día intercalado (Parker 1950).

Un año de exactamente 365 días queda corto en aproximadamente un cuarto de día, de modo que también deriva respecto de las estaciones, solo que mucho más despacio. El sistema egipcio toleró esa deriva y siguió aparte la salida heliaca de Sirio; Parker registra esa salida entre el 17 y el 19 de julio del calendario juliano a lo largo de toda la historia dinástica.

2.3 Conservar ambos, pagar en intercalación

La tercera familia conserva meses lunares verdaderos y los fuerza a reconciliarse con las estaciones insertando un mes adicional a intervalos. La aritmética que lo hace viable es una casi coincidencia:

- Diecinueve años trópicos suman 6939,60161 días.
- Doscientos treinta y cinco meses sinódicos suman 6939,68865 días.
- **Difieren en 0,08704 días, unas 2,09 horas.**

Diecinueve años contienen por tanto casi exactamente 235 meses lunares, es decir doce años ordinarios de doce meses más siete años con un decimotercero. El ciclo de diecinueve años estaba en uso en el calendario babilónico desde comienzos del siglo V a.C., y la cuestión de si el conocimiento griego procedía de Babilonia o fue un hallazgo independiente sigue abierta, citándose un uso chino independiente del mismo ciclo como razón para no descartar el descubrimiento independiente (Steele 2020).

El residuo es pequeño pero no nulo. A unas dos horas por ciclo, acumula unos 0,44 días, o 10,4 horas, tras cinco ciclos, es decir, noventa y cinco años. La práctica calendárica real debe acomodar además que las propias medias astronómicas varían.

En qué punto estamos

Conserva el mes y el año recorre las estaciones, que es lo que hace el islámico por diseño; conserva el año y el civil egipcio deriva un cuarto de día cada vez; consérvalos y el precio es un decimotercer mes, porque diecinueve años y 235 meses difieren en 0,08704 días. Si el uso griego vino de Babilonia sigue abierto, y el ajuste de diecinueve años es próximo, no exacto.

3. El error juliano y qué decidió realmente la reforma gregoriana

Comparado con la media de la época 2000 usada aquí, un año de exactamente 365,25 días excede al año trópico en 0,00781 días. **Como cálculo ilustrativo, equivale a un día cada 128 años**, y a 9,82 días a lo largo de 1.257 años. No es una reconstrucción que suponga constante el año trópico histórico.

Richards informa de que en el siglo XIII ya se reconocía el desplazamiento del equinoccio vernal respecto del 21 de marzo y de que el debate y los intentos de corrección precedieron durante siglos a la reforma de 1582 (Richards 2013). La reforma cambió tanto las reglas de los bisiestos como el método para calcular la Pascua.

El año gregoriano de 365,2425 días, obtenido suprimiendo tres bisiestos cada cuatro siglos, excede al año trópico en 0,00031 días, **lo que por la misma aritmética es un día cada unos 3226 años.**

La reforma no hizo exacto el calendario; redujo el residuo solar a largo plazo. También revisó las tablas lunares eclesiásticas usadas para calcular la Pascua, por lo que no debe describirse simplemente como una elección de las estaciones que ignoró la Luna. La comparación con la época 2000 es una aproximación transparente, no una tasa de error permanente.

En qué punto estamos

Frente a la media de la época 2000 un año juliano excede al año trópico en 0,00781 días y uno gregoriano en 0,00031: un día cada 128 años frente a uno cada 3226. Ambas son ilustrativas: los periodos cambian lentamente. La reforma de 1582 redujo el residuo sin eliminarlo y revisó las tablas lunares eclesiásticas, y no eligió solo el Sol. No todo ciclo apunta a un periodo.

4. Los ciclos calendáricos no tienen por qué ser periodos astronómicos

El supuesto de que un calendario existe solo para seguir un año agrícola o astronómico es demasiado estrecho. Richards señala que los ciclos centroamericanos de trece y veinte días no tienen base astronómica directa (Richards 2013). Su combinación produce una cuenta de 260 días, un periodo calendárico que no es ni un mes sinódico ni un año trópico.

Šprajc, Inomata y Aveni infieren de la orientación de complejos ceremoniales formativos que el calendario de 260 días se usaba entre 1100 y 750 a.C., siglos antes de su registro textual conocido hasta entonces (Šprajc, Inomata y Aveni 2023). Es evidencia revisada por pares de un uso temprano, pero sigue siendo una inferencia arqueológica basada en orientaciones y no un calendario escrito conservado.

El caso establece el punto más estrecho que aquí se necesita: la organización calendárica puede combinar observaciones astronómicas con ciclos numéricos convencionales. El origen y el conjunto completo de funciones de la cuenta de 260 días requieren evidencia especializada; este artículo no les asigna una causa exclusiva.

En qué punto estamos

Los ciclos de trece y veinte días no tienen base astronómica directa, y su combinación da una cuenta de 260 días que no es ni un mes ni un año. Las orientaciones de complejos formativos se leen como evidencia de su uso entre 1100 y 750 a.C., inferencia arqueológica y no calendario escrito conservado. El orden calendárico puede ser convencional, y la unidad misma tiene su propia historia.

5. De la Tierra que gira al átomo de cesio

Antes de 1960 el segundo se definía como $1/86.400$ del día solar medio. En 1960 se redefinió mediante una fracción del año trópico 1900 y, en 1967, la CGPM adoptó una definición atómica basada en el cesio. La formulación de 2018, efectiva con el SI revisado en 2019, define el segundo fijando en 9.192.631.770 Hz el valor numérico de la frecuencia de la transición hiperfina no perturbada del estado fundamental del cesio 133 (BIPM, «Historical perspective: unit of time, second»; BIPM 2019).

La unidad SI ya no se define, por tanto, mediante la rotación terrestre. Esto no significa que toda medida del tiempo sea independiente de la Tierra: el UTC tiene el ritmo del tiempo atómico, pero se ha mantenido próximo al tiempo rotacional, UT1, mediante segundos intercalares.

En 2022 la Conferencia General de Pesas y Medidas adoptó la Resolución 4 de su vigesimoséptima reunión, sobre el uso y el desarrollo futuro del UTC. **La formulación operativa de la propia resolución es que el valor máximo de la diferencia UT1 menos UTC se aumentará en 2035 o antes**, con el problema enunciado como el riesgo de que las discontinuidades por segundo intercalar causen fallos graves en infraestructura digital crítica, y con el encargo de proponer una nueva tolerancia que garantice la continuidad del UTC durante al menos un siglo.

La redacción importa. La resolución decide que se aumentará la divergencia permitida; no fija por sí misma el nuevo máximo ni dice literalmente que se hayan abolido los segundos intercalares.

El caso está relacionado con la sección 1, pero no es idéntico. Desde 1967 la unidad SI es atómica, mientras el tiempo civil conserva una relación convencional con la rotación terrestre. La decisión de 2022 cambia la tolerancia permitida entre esas escalas; no cambia los periodos astronómicos de la sección 1.

En qué punto estamos

El segundo se definía por el día solar medio, luego por una fracción del año trópico de 1900 y desde 1967 por una transición del cesio fijada en 9.192.631.770 Hz. La unidad dejó la rotación terrestre; el UTC no, próximo al UT1. En 2022 la CGPM decidió que esa diferencia se aumentará, sin decir en cuánto. Cuánto se remonta el cómputo deliberado es otra cuestión.

6. Los candidatos más antiguos están en disputa

Una alineación de fosas en Warren Field, Crathes, en Aberdeenshire, fue interpretada por sus investigadores como posible contador de tiempo lunisolar del octavo milenio a.C. y propuesta como evidencia muy temprana de cómputo deliberado del tiempo (Gaffney et al. 2013). El propio artículo usa repetidamente formulaciones cautas como «parece» y «puede».

Más atrás aún, se argumentó que marcas incisas en hueso y asta del Paleolítico superior constituyen una notación lunar secuencial deliberada (Marshack 1972), interpretación discutida en un trabajo cuyo título plantea la objeción como pregunta sobre un deseo tomado por realidad (d'Errico 1989).

La advertencia metodológica de KNO-006 se aplica, pero no sin cambios. En Warren Field el argumento combina la secuencia de fosas, una interpretación lunar y una alineación invernal; en los huesos incisos la inferencia depende más de la organización de las marcas. En ambos casos la función se infiere y no viene escrita en el artefacto. Una secuencia de muescas puede ser compatible con un recuento lunar, con otras cuentas o con patrones no notacionales, por lo que la compatibilidad no establece por sí sola el uso.

Conviene enunciar una limitación concreta de esta ronda. No se localizó ninguna crítica académica publicada que aborde específicamente la interpretación de Warren Field. Es una afirmación sobre esta búsqueda y no evidencia de que la interpretación no esté discutida.

En qué punto estamos

Warren Field fue interpretado por sus investigadores como posible contador lunisolar, con formulaciones que mantuvieron cautas, y no se localizó crítica publicada de esa interpretación, lo que es una afirmación sobre la búsqueda. De unas marcas incisas paleolíticas se argumentó que eran notación lunar deliberada, y el argumento fue discutido. La función se infiere en ambos casos, y para qué servía es aún más difícil.

7. Para qué sirve realmente un calendario

La sección 1 presentó los calendarios como soluciones a una restricción astronómica, lo cual es cierto e incompleto. Un calendario es además un dispositivo de coordinación, lo que ayuda a explicar la necesidad de reglas previsibles, procedimientos compartidos y fechas comunicables.

Los calendarios pueden coordinar caza, agricultura, rituales, contratos y vida social (Richards 2013). Los diseños intercambian previsibilidad por observación: los calendarios aritméticos pueden calcularse mediante reglas fijas, mientras los observacionales esperan una creciente, un solsticio u otro fenómeno y no siempre pueden construirse por completo de antemano. La coordinación exige procedimientos compartidos y una forma de comunicar fechas, no que cada usuario realice personalmente el cálculo.

Esto ayuda a entender por qué una reforma no es solo un cálculo astronómico. Cambiar un calendario también modifica un estándar de coordinación, observancias religiosas y prácticas administrativas. Afirmar que esos costes sociales explican una demora histórica concreta es una interpretación que exige evidencia específica del caso.

En qué punto estamos

El planteamiento de la sección 1 es cierto e incompleto. Un calendario coordina caza, agricultura, ritual, contratos y vida social, así que su diseño tiene consecuencias prácticas de coordinación además de astronómicas. Los diseños aritméticos pueden calcularse de antemano mientras los observacionales esperan una creciente o un solsticio, y coordinar exige procedimiento compartido antes que cálculo universal. Explicar una demora por su coste social pediría evidencia de ese caso.

8. Lo que la evidencia no puede resolver

Si un artefacto temprano es un calendario. Sin evidencia independiente de uso, hacer corresponder un artefacto con una función calendárica moderna es inferir por semejanza.

La intención tras un periodo observado. Que una estructura se alinee con un solsticio no establece que la alineación fuera intencionada, y el número de alineaciones posibles en cualquier estructura es grande.

Si los sistemas tempranos se entendían como aproximaciones. Los usuarios posteriores de un ciclo de diecinueve años sabían que era inexacto. Si lo sabían los primeros no se recupera del uso del ciclo por sí solo.

Cuatro afirmaciones descriptivas dependen de fuentes cuyo texto pertinente no pudo comprobarse directamente; por ello se mantienen deliberadamente limitadas.

En qué punto estamos

La evidencia se detiene en cuatro puntos. Llamar calendario a un artefacto temprano es inferir por semejanza, sin evidencia independiente de uso. Una alineación que existe no fue por fuerza intencionada, porque las alineaciones posibles en cualquier estructura son muchas. Si los primeros usuarios de un ciclo inexacto lo sabían no se recupera del ciclo. Y cuatro afirmaciones descriptivas descansan en fuentes no leídas directamente aquí.

9. Conclusión

Los días medios, los meses lunares y los años trópicos no guardan razones enteras, y los calendarios que intentan seguirlos responden mediante deriva, intercalación, observación o reglas aritméticas. Otros ciclos calendáricos ni siquiera tienen que representar un periodo astronómico.

Leer la historia como progreso simple hacia el calendario gregoriano confunde finalidades distintas con etapas de una sola escalera. Un calendario lunar no es un calendario solar fallido, y una cuenta de 260 días no es un año trópico inexacto. El segundo SI es atómico desde 1967, mientras el UTC sigue mediando entre el ritmo atómico y la rotación terrestre. En 2022 la CGPM decidió que la diferencia UT1-UTC permitida se ampliaría en 2035 o antes; esa resolución no fijó el nuevo máximo.

En qué punto estamos

La escalera es el error: convierte finalidades distintas en etapas y un calendario lunar en uno solar fallido. Días medios, meses lunares y años trópicos no guardan razones enteras, lo que deja elección a los calendarios, y otros ciclos no representan periodo astronómico. Desde 1967 el segundo es atómico y el UTC sigue mediando entre ese ritmo y la Tierra que rota. Qué sabemos, y cómo, es lo que sigue.

Lo que sabemos y cómo lo sabemos

Para la época 2000 una obra astronómica de referencia da un año trópico medio de 365,24219 días y un mes sinódico medio de 29,53059 días, y subraya que ambos son medias aproximadas que cambian lentamente. Todo lo que el artículo deriva de ahí es aritmética propia y así consta: doce meses sinódicos quedan 10,87511 días por debajo de un año, y diecinueve años trópicos difieren de 235 meses sinódicos en 0,08704 días. El año civil egipcio tenía 365 días sin día intercalado, y el calendario islámico mantiene los meses en relación con las lunaciones sin mantener su año con el Sol. Los ciclos centroamericanos de trece y veinte días no tienen base astronómica directa. Y el segundo del SI es atómico desde 1967, fijando su formulación vigente en 9.192.631.770 Hz la transición hiperfina no perturbada del estado fundamental del cesio 133.

Lo que todavía no sabemos

Si Warren Field funcionó como contador lunisolar es una interpretación que sus propios investigadores formularon con cautela, y no se localizó en esta ronda ninguna crítica publicada dirigida específicamente a ella, lo cual dice algo de la búsqueda y no de la literatura. Si las marcas incisas del Paleolítico superior registran lunaciones sigue abierto, porque la compatibilidad con un recuento lunar no lo establece. Si los primeros usuarios de un ciclo de diecinueve años lo entendían como inexacto no

se recupera del uso del ciclo. El origen y las funciones de la cuenta de 260 días exigen evidencia especializada que el artículo no aporta. Y cuatro afirmaciones descriptivas descansan en fuentes cuyo texto pertinente no se leyó directamente.

Qué podría hacernos cambiar de opinión

Evidencia independiente de uso, y no la semejanza con una función calendárica moderna, reforzaría o rompería la lectura de Warren Field, y una crítica publicada dirigida a ella cambiaría lo que el artículo puede decir sobre cuán discutida está. Para los huesos incisivos, un método que separase el recuento lunar de otras cuentas y de patrones no notacionales resolvería lo que las marcas solas no pueden. La lectura directa de las cuatro fuentes delegadas las sacaría del estado parcial. Y la tolerancia que se encargó proponer a la CGPM sustituirá un valor abierto por uno decidido.

Si solo recuerdas una cosa

Un calendario lunar no es un calendario solar fallido: los periodos astronómicos no encajan exactamente, y los calendarios que los siguen deben decidir cómo gestionar el desajuste.

Artículos relacionados

- KNO-006 · *Mapas y representación espacial*
- KNO-008 · *Instrumentos que amplían los sentidos*

Referencias

BIPM. 2019. *SI Brochure, 9th Edition, Appendix 2: Mise en pratique for the Definition of the Second in the SI*. Bureau International des Poids et Mesures, 20 de mayo de 2019. URL estable: <https://www.bipm.org/documents/20126/41489667/SI-App2-second.pdf> **C-D** BIPM. «Historical Perspective: Unit of Time, Second». Bureau International des Poids et Mesures. URL estable: <https://www.bipm.org/en/history-si/second> CGPM. 2022. Resolución 4 de la 27.^a CGPM, «On the Use and Future Development of UTC». Conférence Générale des Poids et Mesures, 15 a 18 de noviembre de 2022. DOI 10.59161/CGPM2022RES4E. URL estable: <https://www.bipm.org/en/-/resolution-cgpm-27-4> **C-D** d'Errico. 1989. «Palaeolithic Lunar Calendars: A Case of Wishful Thinking?». *Current Anthropology* 30 (1). DOI 10.1086/203721. Gaffney, Vincent, Simon Fitch, Eleanor Ramsey, Ron Yorston, Eugene Ch'ng, Eamonn Baldwin, Richard Bates, Christopher Gaffney, Clive Ruggles, Tom Sparrow, Anneley McMillan, Dave Cowley, Shannon Fraser, Charles Murray, Hilary Murray, Emma Hopla, y Andy Howard. 2013. «Time and a Place: A Luni-Solar "Time-Reckoner" from 8th Millennium BC Scotland». *Internet Archaeology* 34. DOI 10.11141/ia.34.1. **C-D** Marshack, Alexander. 1972. *The Roots of Civilization: The Cognitive Beginnings of Man's First Art, Symbol and Notation*. New York: McGraw-Hill. ISBN 978-0-07-040535-6. Parker, Richard A. 1950. *The Calendars of Ancient Egypt*. Studies in Ancient Oriental Civilization 26. Chicago: Oriental Institute of the University of Chicago. URL estable: <https://isac.uchicago.edu/sites/default/files/uploads/shared/docs/saoc26.pdf> Richards, E. G. 2013. «Calendars». Capítulo 15 de *Explanatory Supplement to the Astronomical Almanac*, 3.^a ed., editado por S. E. Urban y P. K. Seidelmann. Mill Valley, CA: University Science Books. URL estable: https://aa.usno.navy.mil/downloads/c15_usb_online.pdf Šprajc, Ivan, Takeshi Inomata, y Anthony F. Aveni. 2023. «Origins of Mesoamerican Astronomy and Calendar: Evidence from the Olmec and Maya Regions». *Science Advances* 9 (1): eabq7675. DOI 10.1126/sciadv.abq7675. Steele, J. M. 2020.

«Geminus and Babylonian Astronomy». En *Instruments, Observations, Theories: Studies in the History of Astronomy in Honor of James Evans*, editado por Alexander Jones y Christián Carman, 205–212. DOI 10.5281/zenodo.3975741.

KNO-008 · Instrumentos que amplían los sentidos

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

La expresión del título de este artículo es una abreviatura útil y una descripción pobre. Un termómetro no afina el sentido del calor de nadie; convierte la temperatura en una longitud que el ojo puede leer. Lo que un instrumento construye es una cadena de conversiones cuyo último eslabón resulta perceptible, y el sensorio humano no se amplía en ningún punto de ella. Este artículo sigue las consecuencias: que fiarse de una lectura es fiarse de todos los eslabones, que calibrar un instrumento es circular antes de ser rutinario, y que los instrumentos interesantes son aquellos cuya salida no corresponde a ningún sentido. Salda el problema de la longitud que dos artículos anteriores aplazaron, y encuentra que el relato habitual sobre cómo se resolvió no sobrevive al contraste con la literatura académica.

En 60 segundos

Unas gafas ayudan a un sentido, un telescopio extiende uno, y un termómetro no hace ninguna de las dos cosas: convierte una magnitud que ningún ojo registra en una longitud que el ojo lee, y el observador infiere. Creer esa lectura es creer todos los eslabones de esa cadena, y ningún punto de ella se zanja mirando. El telescopio de Galileo se aceptó por replicación independiente y no a la vista. La calibración es circular de partida, porque justificar un instrumento exige la ley que medir con él debería establecer; ese círculo se rompió por iteración y no por decreto, y una crítica publicada sostiene que la argumentación razonada puede romperlo. La longitud no la resolvió un reloj: las distancias lunares se trataron como equiparables en 1767, se apoyaban en la estima entre observaciones, y declinaron al abaratare los cronómetros. La radioastronomía empezó como un intento de reducir ruido. Una detección de ondas gravitacionales no termina en ninguna sensación, y su primera señal salió de una búsqueda de ráfagas y no del ajuste de plantillas. Lo que hace comparables dos lecturas es una cadena institucional, y el vocabulario que la define advierte qué no garantiza.

Por qué importa esta pregunta

Una lectura se ofrece como si fuera una observación, y es el final de una cadena que empezó donde ningún sentido llega. Tomarla por percepción esconde todos los lugares donde la cadena puede equivocarse: la sustancia acoplada a la magnitud, la calibración que hay detrás, el modelo que convierte una correlación en una afirmación. Las distinciones de aquí permiten preguntar qué validó el instrumento, si el acuerdo entre copias vale algo cuando comparten diseño, y si el relato de cómo se resolvió un problema es el que el registro sostiene de verdad.

1. Qué hace realmente un instrumento

Conviene distinguir tres dispositivos que suelen archivarse juntos.

Unas gafas corrigen un defecto de un sentido existente y no cambian nada sobre qué clase de cosas puede percibirse. La luz sigue llegando al ojo y el ojo sigue viéndola.

Un telescopio opera en la misma modalidad y empuja más allá de un límite natural. Sigue siendo luz, sigue siendo el ojo, sigue siendo ver, y los objetos que revela son de una clase que el ojo desnudo ya conoce.

Un termómetro hace algo categóricamente distinto. La temperatura no es algo que los ojos perciban. El instrumento acopla una sustancia termométrica a la magnitud de interés y la convierte en una longitud, y la longitud es lo que el observador ve realmente. El observador no percibe la temperatura con más agudeza. El observador percibe una columna de líquido e infiere.

De los tres, las gafas **ayudan** a un sentido y el telescopio **extiende** uno más allá de su rango natural, que son los dos casos literales y los menos interesantes; la tabla de abajo reserva cada palabra para uno de ellos. El tercero es lo que hacen muchos instrumentos científicos, y el patrón se repite ampliamente: el dispositivo convierte una magnitud inaccesible en una accesible, y cada paso de esa conversión es un lugar donde la cadena puede equivocarse.

Esto sugiere un vocabulario de trabajo, ofrecido como estipulación y no como hallazgo.

||| |---|---| | **Ayuda** | Restaura o mejora un sentido dentro de su propia modalidad. | | **Extensión** |
Misma modalidad, más allá del alcance natural. | | **Transductor** | Convierte una magnitud que los
sentidos no registran en otra que sí. | | **Instrumento inferencial** | No produce ningún análogo
perceptible; la salida es un número o una afirmación estadística, y al fenómeno solo se llega a través
de un modelo. |

Los cuatro no son etapas. Coexisten, y un mismo aparato puede contener los cuatro en serie.

En qué punto estamos

Tres dispositivos archivados juntos hacen tres cosas distintas, y solo dos tienen que ver con un sentido. Devolver la vista y alcanzar más allá del rango natural son los casos literales y los menos interesantes. El tercero convierte lo que ningún ojo registra en algo que un ojo sí, y deja al observador como lector de columnas y no como receptor de calor.

2. La cadena hay que crearla eslabón a eslabón

Si un instrumento es una cadena de conversiones, entonces la cuestión de si crearlo no es una cuestión sobre los sentidos. Es una cuestión sobre la cadena, y no hay punto alguno en ella donde la percepción zanje el asunto.

El telescopio es un ejemplo temprano de observación mediada por instrumentos. Las observaciones de Galileo de 1609 y 1610 no se aceptaron a simple vista; el instrumento y sus informes se difundieron por Europa y fueron retomados por múltiples observadores. La secuencia precisa de aquella recepción disputada no se reconstruye aquí, porque las fuentes accesibles no sostienen una cronología más detallada.

Conviene nombrar la estructura porque reaparece. Un instrumento nuevo informa de algo que nadie ha visto. El informe no puede comprobarse mirando, porque mirar es justamente lo que está en cuestión. Una comprobación que funciona es el acuerdo entre instancias independientes de la misma cadena, construidas por gente sin interés en los resultados de la otra. **Por sí sola no basta.** Las copias de un instrumento pueden compartir diseño, ruta de calibración o modelo de fondo, y los instrumentos que comparten cualquiera de esas cosas pueden coincidir por el mismo motivo equivocado. Lo que eleva ese acuerdo son comprobaciones de otra clase: alimentar la cadena con una **entrada conocida** y ver si la devuelve, medir la misma magnitud con un **instrumento de otro diseño**, **intervenir** sobre el fenómeno y ver si la lectura acompaña, y contrastar el resultado con una **restricción independiente** que no venga por la misma cadena.

En qué punto estamos

Si creer una lectura es creer una cadena, mirar más no lo zanja, y el caso fundacional lo confirma: el instrumento nuevo se creyó cuando gente sin interés en los resultados ajenos construyó el suyo y obtuvo lo mismo. El acuerdo por sí solo no basta, porque unas copias pueden compartir diseño y equivocarse juntas. Entradas conocidas, otros diseños, intervención y restricciones externas son lo que lo eleva.

3. La calibración es circular antes de ser rutinaria

El problema más hondo de la instrumentación no es construir el dispositivo. Es establecer qué significan sus lecturas, y ese problema es circular de partida.

Para justificar un termómetro hace falta la ley que relaciona el comportamiento de la sustancia termométrica con la temperatura. Para establecer esa ley hace falta medir la temperatura. La circularidad es real y tiene nombre: el problema de la medición nómica, expuesto en un estudio de la termometría de los siglos XVIII y XIX que sostiene que el círculo no se rompió por decreto sino por refinamiento iterativo y autocorrectivo entre tradiciones termométricas rivales a lo largo de un periodo largo (Chang 2004).

De ahí se siguen dos consecuencias y las dos se pasan por alto con facilidad.

Primera, **un instrumento no se limita a medir una magnitud que ya estuviera bien definida.** Antes de una termometría practicable no había manera operativa de decir que dos cuerpos estaban a la misma temperatura, y fijar el procedimiento de medida fue parte de fijar qué era la magnitud.

Segunda, **una circularidad de esta clase puede aparecer donde esté en juego un fenómeno nuevo**, y es un riesgo que conviene buscar y no la estructura de todo instrumento nuevo. Donde aparece, si un aparato funciona correctamente se juzga por si produce el resultado esperado, y qué resultado esperar es lo que el aparato se construyó para averiguar. **El caso de la medición nómica del párrafo anterior es aún más estrecho:** pertenece a los casos fundacionales en que la propia magnitud se está fijando junto con su medida, y un termómetro contrastado con un patrón mantenido no está en esa situación. En sociología de la ciencia esto se llama regreso del experimentador y se asocia a Collins (1992). La edición de University of Chicago Press de 1992 incluye *Detecting Gravitational Radiation: The Experimenter's Regress* como capítulo 4. Ese dato bibliográfico sostiene la atribución, pero no todas las interpretaciones más amplias del argumento del capítulo.

Hay una crítica publicada del regreso y su sitio es este y no una nota al pie. Franklin sostiene, a partir del episodio de las ondas gravitacionales, que el regreso se rompió mediante argumentación razonada (Franklin 1994). Léase con la estrechez con que está escrito: niega que el regreso sea

ineludible, no que nunca aparezca. Esta sección ya trata la circularidad como un riesgo que conviene buscar y no como la estructura de todo instrumento nuevo, y la crítica es lo que convierte ese tratamiento en una posición dentro de un debate y no en una cautela.

En qué punto estamos

Establecer qué significa una lectura cuesta más que construir el aparato, y al principio los dos requisitos se persiguen. La historia muestra el círculo desgastado por tradiciones rivales que se corrigen entre sí y no cortado por decreto. Fijar cómo medir fue parte de fijar qué se medía, y la circularidad es un riesgo que rastrear, no la forma de todo aparato nuevo. Una crítica publicada sostiene que puede romperse.

4. La longitud, y una deuda saldada

KNO-006 aplazó el problema de la longitud por corresponder al tiempo. KNO-007 lo examinó y lo remitió aquí, por no ser una cuestión sobre qué es el tiempo sino sobre transportar un patrón horario. Este artículo salda la deuda, y la literatura académica desmonta el relato habitual.

La versión popular dice que los cronómetros marinos de John Harrison resolvieron el problema. La versión académica dice que hallar la longitud en la práctica exigía instrumentos, procedimientos y tablas, y personas con criterio, y que ningún artefacto aislado resolvió nada por sí solo (Miller 2015).

Más llamativo es el destino del método rival. Las distancias lunares, hechas practicables por las tablas de Tobias Mayer en la década de 1750 y por el *Nautical Almanac* de Maskelyne desde 1767, fueron tratadas por el Board of Longitude en 1767 como equiparables a los cronómetros. El viaje del *Endeavour* de 1768 a 1771 usó distancias lunares, en un momento en que el H4 de Harrison era el único cronómetro suficientemente exacto y solo existía uno. **Este artículo no dice que fueran el método principal**, ni allí ni en los demás viajes de Cook, porque es una afirmación que las páginas accesibles no zanján. Las tablas lunares siguieron en los almanaques náuticos bien entrado el siglo siguiente, y el método declinó a medida que los cronómetros se abarataron, no porque fallara (de Grijns 2020).

Esa misma fuente describe además la estima en que el método se apoyaba. En el viaje del *Endeavour* el trabajo lunar se combinó con la estima, contrastada y probada, que siguió siendo una herramienta de navegación fundamental aunque supletoria cuando el tiempo interrumpía las observaciones lunares; la velocidad del barco se estimaba con una barquilla en una corredera de nudos cronometrada con una ampolleta, y el procedimiento no podía corregir la deriva ni las corrientes (de Grijns 2020). **Aquí no se da ninguna cifra de longitud**, porque lo que el pasaje aporta es un procedimiento y sus límites y no un número.

Lo general de esto va sobre instrumentos y no sobre barcos. **Una solución instrumental y una solución computacional pueden ser igualmente correctas**, y cuál se impone puede decidirlo el coste de fabricación y no la exactitud. Leer el desenlace hacia atrás como demostración de que el reloj era la respuesta correcta importa una conclusión que la historia no proporciona.

En qué punto estamos

Una deuda que dos artículos anteriores dejaron abierta se cierra aquí, y la respuesta habitual resulta comprimir una historia más larga. Hallar la posición en el mar exigió a la vez aparatos, tablas, procedimientos y criterio, sin que ningún artefacto decidiera nada solo. El método rival se trató como igual, se apoyó en una estima ciega a las corrientes, y se apagó cuando su competidor se abarató.

5. Instrumentos sin análogo perceptivo

Algunos instrumentos detectan cosas para las que no podría existir un sentido, y son los que muestran la estructura de transducción con mayor claridad.

La radioastronomía empezó por accidente y no por diseño. Karl Jansky estaba empleado en rastrear la estática que interfería en la telefonía transatlántica (NRAO/AUI Archives, «Papers of Karl G. Jansky»), con una antena rotatoria (Jansky 1932) que operaba a 21 MHz (Kellermann et al. 2023). Primero sospechó del Sol. A comienzos de 1933 la dirección del siseo ya no seguía al Sol, y en febrero concluyó que procedía de una dirección fija en el espacio (Kellermann et al. 2023), hacia el centro de la Vía Láctea (NRAO, Jansky Lectureship). El artículo primario registra que se había pensado que la fuente de esa perturbación estaba asociada de algún modo al Sol (Jansky 1933), que es la hipótesis que el análisis posterior desplazó, y esa misma historia institucional identifica el resultado como ondas de radio procedentes de la región central de nuestra Vía Láctea. El instrumento no se construyó para observar nada. Se construyó para reducir ruido, y el ruido era la observación.

La detección de ondas gravitacionales no tiene ninguna etapa perceptible. La primera observación directa comunicada da una deformación máxima de $1,0 \times 10^{-21}$ (Abbott et al. 2016). Nada en esa cadena termina en una sensación. Además la detección misma es un juicio estadístico. **Uno de los métodos que se usan para ella** es el filtrado adaptado: los datos se correlacionan con un banco de formas de onda teóricas, y para ruido gaussiano eso se reduce a una correlación maximizada entre datos y plantilla que optimiza la relación señal a ruido, que es una forma de detección por máxima verosimilitud (Jaranowski y Królak 2012). **Ese método no agota el modo en que se halla una señal así**, y este artículo no lo presenta como la definición de la detección.

La detección inicial no salió del ajuste de plantillas en absoluto. El propio resumen de la colaboración registra que la detección inicial la hizo una búsqueda llamada de ráfagas, y que dos tuberías separadas de filtrado adaptado, llamadas PyCBC y GstLAL, recuperaron el suceso después (LIGO Scientific Collaboration, resumen científico). De modo que las plantillas fueron una comprobación cruzada sobre una detección hecha sin ellas, y un relato que empieza por el filtrado adaptado tiene el orden del revés.

Conviene decirlo sin rodeos. **El objeto de la afirmación no se ve, no se oye y no lo registra ningún sentido. Se infiere del grado en que un registro se ajusta a un modelo de lo que produciría tal suceso.** El instrumento no amplía un sentido. Sustituye la percepción por inferencia, y allí donde la inferencia pasa por plantillas vale lo que valga el modelo que las suministra.

En qué punto estamos

Algunos instrumentos alcanzan objetos que ningún sentido registra, y ahí la estructura se ve clara. Uno empezó limpiando interferencias y acabó diciendo de dónde venía un siseo. Otro no termina en sensación ninguna, y llega a su objeto mediante un modelo de lo que lo produciría. Incluso allí la señal inicial se halló sin ajuste de plantillas, y las tuberías de filtrado adaptado la recuperaron después.

6. Qué hace comparables dos lecturas

Un instrumento en el que solo confía su dueño sirve de poco. Lo que hace comparable una lectura de un laboratorio con la de otro es la trazabilidad metrológica, la cadena de calibración hasta una referencia común, concepto incorporado a la tercera edición del vocabulario internacional de metrología (JCGM 200:2012).

Aquí corresponde una cautela de cita y no una nota al pie. **Una copia muy difundida de ese vocabulario lleva la fecha de 2006 y es un borrador de comité previo a la publicación, no la edición publicada.** Quien la citara desde esa copia estaría citando un borrador. La edición de referencia vigente es JCGM 200:2012.

El vocabulario la define como propiedad de un resultado, y dice qué no garantiza. La trazabilidad metrológica es una propiedad de un resultado de medida por la cual el resultado puede relacionarse con una referencia mediante una cadena documentada e ininterrumpida de calibraciones, cada una de las cuales contribuye a la incertidumbre de medida; y la misma entrada advierte que la trazabilidad no asegura que la incertidumbre de medida sea adecuada para un fin dado, ni asegura la ausencia de errores (JCGM 200:2012, VIM 2.41, Note 5).

El punto conceptual es que la comparabilidad no es una propiedad de un instrumento. Es una propiedad de un arreglo institucional que enlaza unos instrumentos con otros, y hay que mantenerlo. Un instrumento sin calibrar no es un instrumento algo peor; es un instrumento cuyas lecturas no conectan con las de nadie.

En qué punto estamos

Un instrumento que nadie más puede comprobar sirve de poco, y lo que conecta un laboratorio con otro es una cadena mantenida y no una propiedad de ningún aparato. El vocabulario que define esa cadena es cuidadoso con lo que no promete: ni que la incertidumbre convenga al fin, ni que no se haya cometido ningún error. Un instrumento sin calibrar no es algo peor; sus lecturas no conectan.

7. Lo que la evidencia no puede resolver

Si un instrumento funciona, en un caso nuevo. La circularidad de la sección 3 no es un acertijo a la espera de solución. Es un rasgo estructural, gestionado por iteración y por replicación independiente en vez de eliminado.

Qué habría mostrado un instrumento construido de otro modo. Todo instrumento selecciona. Lo que no registra no está por ello ausente, y la ausencia de señal es evidencia solo cuando se conoce la respuesta del instrumento a esa señal.

La prioridad en el descubrimiento accidental. Cuando un instrumento construido para un fin registra otra cosa, decidir quién descubrió qué depende de qué cuenta como reconocer un fenómeno, y eso no lo zanja el registro por sí solo.

Tres obras citadas pudieron verificarse bibliográficamente, pero no leerse por completo en los pasajes pertinentes; las afirmaciones que dependen de ellas se limitan por ello a lo que pudo establecerse.

En qué punto estamos

Cuatro cosas quedan abiertas, cada una por un motivo distinto. Si un aparato funciona en un caso realmente nuevo es estructural y no un problema sin resolver. El contrafáctico sobre otro diseño carece de respuesta, porque seleccionar es lo que los aparatos hacen. El crédito por lo no buscado depende de qué cuenta como reconocer. Y tres fuentes se verificaron como publicaciones sin leer su texto pertinente.

8. Conclusión

Muchos instrumentos no amplían los sentidos en absoluto. Los alistan al final de cadenas que empiezan donde los sentidos no llegan, y en esos casos el peso epistémico recae en la cadena y no en el momento de percepción que la remata. **Algunos instrumentos sí amplían un sentido**, y la sección 1 los separa; lo que este artículo sostiene vale para los muchos, no para todos.

Por eso la historia de los instrumentos no es una historia de ojos más agudos. Es una historia de calibración, de replicación entre gente que no se fía entre sí, de patrones sostenidos por instituciones, y de modelos que convierten una correlación en una afirmación sobre la fusión de dos agujeros negros. Cuando se cree a un instrumento, lo que se está creyendo es todo eso, y la lectura es lo último y lo menos.

En qué punto estamos

El argumento vuelve al lugar donde recae el peso. En muchos instrumentos la carga epistémica la lleva la cadena y no el momento de percepción que la remata, y la sección 1 acota aquellos en que no es así. Así leída, la historia deja de ser la de unos ojos mejores y pasa a ser la de la calibración, la desconfianza y los patrones mantenidos.

Lo que sabemos y cómo lo sabemos

Expertos de Bolonia informaron de que veían poco o nada por el telescopio de Galileo en abril de 1610, el Colegio Romano retuvo la confirmación hasta octubre, y la credibilidad se estableció por replicación independiente; la fuente de esto se leyó solo hasta donde llegan sus páginas accesibles. La circularidad de validar un instrumento de medida sin la ley que lo justifica se llama problema de la medición nómica y se atribuye a un estudio de la termometría de 2004, que sostiene que el círculo se desgastó por refinamiento iterativo y autocorrectivo y no se rompió por decreto. Franklin sostiene, a partir del episodio de las ondas gravitacionales, que la argumentación razonada rompió el regreso. La edición de University of Chicago Press de 1992 titula su capítulo 4 con la expresión *Detecting Gravitational Radiation: The Experimenter's Regress*, que la autoridad humana leyó en el índice y que nadie leyó más allá. El Board of Longitude trató en 1767 los cronómetros y las distancias lunares como equiparables; el viaje del *Endeavour* combinó el trabajo lunar con la estima, que no podía corregir la deriva ni las corrientes. La historia institucional del observatorio identifica el resultado de Jansky como ondas de radio procedentes de la región central de la Vía Láctea, y el artículo primario registra la hipótesis solar que quedó desplazada. La primera observación directa comunicada de ondas gravitacionales da una deformación máxima de $1,0 \times 10^{-21}$, y la detección inicial la hizo una búsqueda de ráfagas, recuperando después el suceso PyCBC y GstLAL. La trazabilidad metrológica se define como propiedad de un resultado de medida ligado a una referencia por una cadena documentada e ininterrumpida de calibraciones.

Lo que todavía no sabemos

Varios detalles quedan fuera del alcance evidencial utilizado aquí: el argumento completo de Collins sobre el regreso del experimentador; la recepción disputada precisa del telescopio de Galileo; las conclusiones detalladas del estudio histórico de la termometría; si las distancias lunares fueron el método principal en los viajes de Cook; la redacción del prólogo publicado en 2012 frente al borrador previo del comité; y si el capítulo de Sullivan respalda de forma independiente el relato sobre Jansky. Ninguno de esos puntos se utiliza aquí como base de una afirmación más fuerte.

Qué podría hacernos cambiar de opinión

Una lectura del capítulo de Collins zanjaría qué sostiene esa obra sobre el regreso del experimentador, frente a que lo nombre. Una fuente cuyo texto pueda leerse, sobre la recepción del telescopio de Galileo, sustituiría una afirmación que hoy descansa en una página de editor y en un ensayo distinto. Evidencia paginada de que las distancias lunares fueron o no el método principal en los viajes de Cook decidiría una cuestión que este artículo deja abierta en ambos sentidos. Leer el prólogo publicado de 2012 convertiría una presunción sobre el vocabulario en una comprobación. Y que un instrumento de diseño enteramente distinto coincidiera con alguno de estos resultados haría más por él que otra copia del mismo diseño coincidiendo, que es lo que sostiene la sección 2 y que vale para las fuentes de este mismo artículo.

Si solo recuerdas una cosa

Un instrumento rara vez amplía un sentido; remata una cadena de conversiones, y creer la lectura es creer todos los eslabones.

Artículos relacionados

- KNO-006 · *Mapas y representación espacial*
- KNO-007 · *Calendarios y medida del tiempo*

Referencias

Este artículo cita en el propio cuerpo, por autor y fecha o por designación de documento, y no lleva números de afirmación entre corchetes. Cada cita resuelve a una fila de

KNO-008_CLAIMS_AND_SOURCES_RECONSTRUCTION.md, que nombra una fuente académica u oficial por afirmación y, desde el 11 de septiembre de 2026, la localización del registro en que se leyó. Las afirmaciones inciertas, los límites de alcance declarados y las fuentes excluidas están en KNO-008_UNCERTAINTIES_RECONSTRUCTION.md.

Abbott, B. P., et al. (LIGO Scientific Collaboration and Virgo Collaboration). 2016. «Observation of Gravitational Waves from a Binary Black Hole Merger». *Physical Review Letters* 116 (6): 061102. DOI 10.1103/PhysRevLett.116.061102.

Bucciantini, Massimo, Michele Camerota, y Franco Giudice. 2015. *Galileo's Telescope: A European Story*. Trad. Catherine Bolton. Cambridge, MA: Harvard University Press. ISBN 978-0-674-73691-7.

Chang, Hasok. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford Studies in Philosophy of Science. Oxford: Oxford University Press. ISBN 978-0-19-517127-3.

Collins, Harry. 1992. *Changing Order: Replication and Induction in Scientific Practice*. Chicago: University of Chicago Press. ISBN 978-0-226-11376-0.

de Grijs, Richard. 2020. «A (not so) brief history of lunar distances: Lunar longitude determination at sea before the chronometer». *Journal of Astronomical History and Heritage* 23 (3): 495–522.

Franklin, Allan. 1994. «How to Avoid the Experimenters' Regress». *Studies in History and Philosophy of Science Part A* 25 (3): 463–491. DOI 10.1016/0039-3681(94)90062-0

Jansky, Karl G. 1932. «Directional Studies of Atmospherics at High Frequencies». *Proceedings of the Institute of Radio Engineers* 20 (12): 1920–1932.

Jansky, Karl G. 1933. «Radio Waves from Outside the Solar System». *Nature* 132: 66. Publicado el 8 de julio de 1933. DOI 10.1038/132066a0

Jaranowski, Piotr, y Andrzej Królak. 2012. «Gravitational-Wave Data Analysis. Formalism and Sample Applications: The Gaussian Case». *Living Reviews in Relativity* 15 (1): art. 4. DOI 10.12942/lrr-2012-4.

JCGM. 2012. *International Vocabulary of Metrology: Basic and General Concepts and Associated Terms (VIM)*. 3.^a edición, JCGM 200:2012, versión corregida. Bureau International des Poids et Mesures. URL estable: <https://www.bipm.org/en/committees/jc/jcgm/wg/jcgm-wg2-vim>

Kellermann, Ken, Ellen Bouton, Heather Cole y Jeff Hellerman. 2023. «Jansky @90: The Origins of a New Window on the Universe». National Radio Astronomy Observatory, 27 de abril de 2023. URL estable: <https://public.nrao.edu/news/jansky-90-the-origins-of-a-new-window-on-the-universe/>

LIGO Scientific Collaboration. Resumen científico «How we searched for merging black holes and found GW150914». La página estable consultada no muestra fecha de publicación; no se asigna una fecha no documentada. URL estable: <https://ligo.org/science-summaries/gw150914cbc/>

Miller, David Philip. 2015. «Longitude Networks on Land and Sea: The East India Company and Longitude Measurement "in the Wild", 1770–1840». En *Navigational Enterprises in Europe and its Empires, 1730–1850*, ed. Richard Dunn y Rebekah Higgitt, 223–247. Londres: Palgrave Macmillan. ISBN 978-1-137-52063-0. DOI 10.1057/9781137520647_12.

National Radio Astronomy Observatory. Página «Jansky Lectureship». La página estable consultada no muestra fecha de publicación; no se asigna una fecha no documentada. URL estable: <https://science.nrao.edu/science/jansky-lecture/>

National Radio Astronomy Observatory, NRAO/AUI Archives. «Papers of Karl G. Jansky, 1923–1998». Inventario del fondo. URL estable: <https://www.nrao.edu/archives/jansky-finding-aid>

Sullivan, W. T., III. 1984. «Karl Jansky and the discovery of extraterrestrial radio waves». En *The Early Years of Radio Astronomy: Reflections Fifty Years after Jansky's Discovery*, 3–42. Cambridge: Cambridge University Press. DOI 10.1017/CBO9780511564956.003.

Nota de frontera

Con KNO-007, Calendarios y medida del tiempo. KNO-007 trata qué se acuerda medir; este artículo trata los dispositivos y las cadenas que lo realizan. **El problema de la longitud, aplazado por KNO-006 y remitido por KNO-007, queda saldado aquí en la sección 4 y no sigue abierto en ningún otro punto de la serie.**

Con KNO-011, Medir antes de teorizar. Es la frontera más próxima de la serie y exige cuidado. **KNO-008 cubre la circularidad de validar un instrumento concreto**, porque es un problema sobre el dispositivo. **KNO-011 cubre la medición sin teoría como estrategia epistémica**, incluida la cuestión de si un programa de investigación puede avanzar midiendo primero, y esa cuestión más amplia no se anticipa aquí.

Con KNO-013, Modelos, mapas y realidad. La sección 5 afirma que la detección de ondas gravitacionales alcanza su objeto a través de un modelo. **La cuestión general de qué es un modelo y cómo afecta la dependencia de modelo al realismo pertenece a KNO-013** y no se desarrolla aquí.

Con KNO-006. El parecido estructural entre la selectividad de un instrumento y la de un mapa es real y limitado, y este artículo no lo fuerza. La proyección cartográfica está restringida por un teorema geométrico; la selectividad instrumental es una decisión de diseño. **Son análogas en su consecuencia y distintas en su naturaleza.**

KNO-009 · Archivos, bibliotecas y memoria colectiva

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

Un archivo no es un contenedor del pasado. Es el resultado de una larga serie de decisiones sobre qué guardar, y esas decisiones no sobreviven en el material que producen, así que un archivo es evidencia sobre el guardar tanto como sobre lo guardado, y sus fondos supervivientes no bastan para revelar la forma completa de lo que falta. KNO-009 separa documentos, archivos y bibliotecas por la lógica de selección de cada uno, toma al pie de la letra la definición profesional de la valoración como decisión sobre qué destruir, y usa un relato en cuatro momentos de dónde entran los silencios. Examina el archivo como instrumento de gobierno, Alejandría como caso donde la explicación dramática es la más débil, y por qué lo digital exige acción continua para persistir.

En 60 segundos

Documentos, archivos y bibliotecas se agrupan y funcionan con principios incompatibles, así que cada uno calla algo distinto: el archivo sobre quienes no generaron documentos, la biblioteca sobre lo que nunca entró en sus circuitos, el documento sobre lo que su finalidad no exigía. Lo que entra en un archivo lo decide la valoración, cuyo planteamiento profesional es qué conservar y qué destruir, y la serie superviviente se lee como completa porque la destrucción no deja hueco en ella. Trouillot separa cuatro momentos donde entra el silencio, y eso importa porque son cuatro fallos con cuatro remedios y buscar mejor solo arregla algunos. La descripción gobierna después lo recuperable, que es una restricción y no una segunda valoración, y puede rehacerse donde la valoración no. Alejandría es el caso donde el deterioro ordinario y la financiación institucional retirada bastan para explicar la desaparición sin acto hostil, que es el argumento de Bagnall y aquí se da como suyo. El caso digital repite el mismo punto estructural en forma moderna: una colección no es una cosa que perdura salvo que la destruyan, sino un proceso que se detiene si no se mantiene, y el gasto pasa a ser continuo.

Por qué importa esta pregunta

Tratar un archivo como contenedor convierte sus fondos en una muestra del pasado y sus huecos en accidente. Las dos cosas son falsas de un modo que cambia lo que puede concluirse. Los fondos son el resultado de un filtro que no se describe a sí mismo, así que solo desde los fondos una ausencia no es evidencia de ausencia ni una supervivencia lo es de tipicidad. Por eso leer bien un archivo no es cuestión de buscar más, y por eso una catástrofe, fácil de narrar, tiende a desplazar una retirada de financiación, que no lo es.

1. Tres instituciones que no son una sola cosa

Documentos, archivos y bibliotecas se tratan habitualmente juntos y funcionan con principios incompatibles.

Un documento se produce para tramitar o documentar una actividad; que se conserve después es otra cuestión. Un recibo existe para saldar una cuenta; un acta existe para fijar una decisión. Su valor probatorio viene de la actividad para la que se creó, mientras que su conservación posterior introduce un filtro distinto, y eso es justamente lo que lo hace buena evidencia y evidencia no representativa a la vez.

Un archivo conserva los documentos de una actividad, y su unidad es el conjunto que generó el organismo productor. Ese productor es a menudo una institución y a menudo una persona o una familia: los papeles personales y las colecciones adquiridas son archivos por el mismo principio, y usar archivo como sinónimo de documentación institucional los deja fuera. Su principio de organización es la procedencia, que es un principio de contexto y de mantener separado el material de orígenes distintos, y no una regla para elegir qué se ingresa. Lo que contiene lo sigue determinando la actividad de quién quedó documentada y los documentos de quién se juzgaron dignos de ingreso, y ese juicio es la valoración, de la que trata por su cuenta la sección 2.

Una biblioteca reúne obras hechas para circular. Su principio de selección es la materia y el uso. Buena parte de lo que guarda fue publicado, lo que significa que esa parte ya pasó por otro filtro antes de que la biblioteca lo viera, y las bibliotecas guardan también material inédito, de modo que biblioteca y publicado no son la misma categoría.

Los tres tienen por tanto silencios distintos. Un archivo calla sobre quienes no generaron documentos o cuyos documentos nadie conservó. Una biblioteca calla sobre lo que nunca entró en los circuitos de los que colecciona. Un documento calla sobre lo que su finalidad no le exigía declarar. **Tratar los tres como un único depósito del pasado funde tres distorsiones distintas en una vista aparentemente única.**

En qué punto estamos

Tres instituciones se agrupan y funcionan con principios incompatibles. Un documento se produce para tramitar o documentar una actividad, y que se conserve después es otra cuestión; un archivo conserva lo que una actividad generó y lo ordena por procedencia; una biblioteca reúne lo hecho para circular. Cada uno es ciego en un sitio distinto, y juntarlos funde tres distorsiones en una vista aparente.

2. La valoración es una decisión sobre qué destruir

La palabra que usa la profesión para decidir qué ingresa en un archivo es valoración, definida como el proceso de identificar los materiales ofrecidos a un archivo que tienen valor suficiente para ser ingresados, con el debate profesional de fondo formulado como qué guardar y qué destruir (SAA Dictionary of Archives Terminology).

Esa formulación merece tomarse al pie de la letra. No se puede conservar todo, de modo que un juicio profesional determina qué dejará de estar disponible a través de esa serie archivística, y el juicio se emite bajo las restricciones y los supuestos de su propio momento. La teoría archivística ha reconocido al archivero como valorador activo y no como conservador pasivo, y ha trazado una secuencia de paradigmas cambiantes en la comprensión de ese papel (Cook 2013).

De ahí se siguen dos consecuencias.

La decisión es irreversible, y la serie superviviente no la muestra. Lo destruido no deja hueco en ella; la serie se lee sin más como completa. Eso es una propiedad de la serie y no del archivo en conjunto: las decisiones de valoración pueden documentarse aparte, en calendarios de conservación e informes de valoración, y donde así ocurre lo irrecuperable es el contenido destruido y no el hecho de que se tomó una decisión.

La composición del archivo es evidencia sobre quien valoró. Lo que una época consideró digno de conservar es un hecho sobre esa época, y es legible en el archivo lo pretendiera alguien o no.

En qué punto estamos

La profesión llama valoración a decidir qué ingresa, y plantea el debate de fondo como qué guardar y qué destruir. El juicio es irreversible y la serie superviviente se lee como completa, aunque los calendarios de eliminación y los informes de valoración dejen constancia de que hubo juicio. Lo que una época guardó es evidencia sobre esa época.

3. Por dónde entran los silencios

El análisis más útil de cómo se producen las ausencias no trata el silencio como un suceso único. Localiza cuatro momentos distintos:

Trouillot los sitúa en la creación del hecho, que es la producción de fuentes; el ensamblaje del hecho, que es la producción de archivos; la recuperación del hecho, que es la producción de relatos; y la significación retrospectiva, que es la producción de historia en última instancia (Trouillot 1995, 26).

El valor de la división en cuatro partes es que separa cosas que suelen confundirse. Puede que una fuente nunca se produjera. Una fuente existente puede no haber ingresado nunca en un archivo. Un documento archivado puede no encontrarse ni usarse nunca. Y el material encontrado puede juzgarse irrelevante. **Son cuatro fallos distintos con cuatro remedios distintos, y ninguna mejora en la búsqueda arregla los dos primeros.**

Explica también por qué un archivo no puede auditarse a sí mismo solo con lo que conserva. El silencio puede entrar en cada momento, y los fondos supervivientes por sí solos no revelan qué momento produjo una ausencia concreta.

En qué punto estamos

Trouillot sitúa el silencio en cuatro momentos: creación, ensamblaje y recuperación del hecho, y significación retrospectiva. El valor está en que son cuatro fallos distintos con cuatro remedios distintos, y buscar mejor no toca los dos primeros. La misma división explica que los fondos supervivientes no puedan auditarse solos: el silencio puede entrar en cada momento, y por sí solos no revelan dónde se produjo una ausencia concreta.

4. El archivo como instrumento de gobierno

Si los archivos se producen en vez de encontrarse, entonces su producción es objeto de estudio por derecho propio. Es el movimiento que suele llamarse giro archivístico: leer los archivos como lugares de producción de conocimiento y como monumentos de los estados, y no como depósitos neutrales de los que extraer hechos (Stoler 2002).

El punto práctico no depende de la teoría. **Un archivo responde a las preguntas que sus creadores necesitaban responder.** Los expedientes de una administración colonial son excelente evidencia sobre lo que preocupaba a esa administración, y mala evidencia sobre la vida de quienes gobernaba, y esa segunda carencia no es un hueco que se llene mirando con más ahínco en los mismos expedientes.

En qué punto estamos

Leer los archivos como lugares de producción de conocimiento y monumentos de los estados, y no como depósitos neutrales, convierte su producción en asunto propio. El punto práctico se sostiene sin la teoría: un archivo responde a las preguntas que sus creadores necesitaban responder, así que los expedientes de una administración colonial son buena evidencia de sus preocupaciones y mala de los gobernados.

5. Alejandría, donde la explicación dramática es la más débil

La destrucción de la Biblioteca de Alejandría es el emblema habitual de la pérdida catastrófica de conocimiento, y suele atribuirse a un incendio.

Un argumento académico apunta a otro sitio. Según Bagnall, el deterioro ordinario del papiro, agravado por la pérdida de voluntad institucional y de financiación para seguir copiando textos, basta para explicar la desaparición sin acto hostil alguno, y él lo formula diciendo que incluso sin acción hostil la Biblioteca, o las Bibliotecas, de Alejandría no habrían sobrevivido a la Antigüedad (Bagnall 2002).

Ese es su argumento y aquí se da como suyo, no como resultado zanjado del campo.

Esto importa más allá de una biblioteca. Una colección sobre soporte perecedero no es algo que persiste salvo que se destruya; es un proceso que se detiene salvo que se mantenga. Copiar es el mecanismo de supervivencia, copiar cuesta dinero e instituciones, y el modo de fallo no es un suceso sino una retirada de atención. **Una catástrofe se narra mejor que un presupuesto, y esa es una de las razones de que persista el relato catastrófico.**

En qué punto estamos

Según Bagnall, el deterioro ordinario del papiro, con la voluntad institucional y la financiación para seguir copiando retiradas, basta para explicar la desaparición sin acto hostil. Es su argumento y se da como suyo. El punto más amplio es que una colección en soporte perecedero es un proceso que se detiene y no una cosa que perdura salvo que la destruyan.

6. Cuando lo que no puede encontrarse desaparece en la práctica

Un archivo que guarda material que nadie puede localizar equivale funcionalmente a uno que no lo guarda. El acceso lo gobierna la descripción, y la descripción sigue normas; la norma internacional general de descripción archivística fue publicada en segunda edición por el Consejo Internacional de Archivos (ISAD(G) 2000).

La consecuencia es que **la descripción gobierna lo que puede recuperarse**, que no es lo mismo que una segunda valoración. Lo descrito a nivel de unidad documental puede encontrarse; lo descrito solo al nivel de un agregado grande solo lo encuentra quien ya sospecha que está ahí. El efecto sobre

quien investiga puede parecerse al de una destrucción, y ambas difieren justo en lo que importa: una ausencia en un instrumento de descripción no demuestra una ausencia en los fondos, y la descripción, a diferencia de la valoración, puede rehacerse.

Aquí corresponde una frontera. Cómo se construyen las categorías, qué hacen a las cosas que ordenan y si una clasificación puede ser neutral son objeto de KNO-010 y no se desarrollan en este artículo. KNO-009 trata la descripción solo como restricción de la recuperación.

En qué punto estamos

La descripción gobierna lo que puede recuperarse, y eso no equivale a una segunda valoración. Lo descrito a nivel de unidad documental se encuentra; lo descrito solo como agregado grande lo encuentra quien ya sospecha. Para quien investiga el efecto puede parecerse a la destrucción, y difiere donde cuenta: un instrumento de descripción no es el fondo, y describir puede rehacerse.

7. La paradoja digital

Suele suponerse que el material digital es permanente porque se copia sin pérdida. El supuesto es falso en su forma general, y también lo es su imagen invertida. El pergamino no sobrevive indefinidamente al abandono, y no todo objeto digital exige migración continua. Lo que se sostiene es más estrecho: un objeto digital depende de que su soporte y su formato sigan siendo legibles, y donde alguno deja de serlo necesita intervención activa y una institución que siga prestándola.

Dos síntomas medidos.

Las referencias a material web se degradan. Un estudio de más de un millón de referencias procedentes de más de tres millones y medio de artículos publicados entre 1997 y 2012 informa de que uno de cada cinco artículos de ciencia, tecnología y medicina sufre degradación de referencias, en el sentido de que resulta imposible volver al contexto web que los rodea algún tiempo después de su publicación (Klein et al. 2014).

La respuesta institucional es un modelo de referencia para sistemas abiertos de información de archivo, vigente como ISO 14721:2025, tercera edición, mantenido conjuntamente con el comité de sistemas de datos espaciales como CCSDS 650.0-M-3.

El punto estructural es el mismo de la sección 5 en forma moderna. La supervivencia no es el estado por defecto de una colección. Es el resultado de un gasto continuo, y el caso digital hace ese gasto continuo en vez de periódico.

En qué punto estamos

Lo digital dura mientras el soporte y el formato admitan lectura, y cuando alguno falla hace falta intervención activa y una institución que siga prestándola. Un estudio sobre más de un millón de referencias cifra en uno de cada cinco los artículos de ciencia, tecnología y medicina con corrosión de referencias. Sobrevivir es gasto continuo.

8. Lo que la evidencia no puede resolver

El tamaño y la forma de lo que falta. Un archivo no puede medir sus propias omisiones, porque la evidencia que las mediría es lo ausente. Las estimaciones de pérdida vienen de fuera del archivo o no vienen.

Si una omisión fue deliberada. Destrucción, abandono, coste y censura pueden dejar el mismo rastro, que es ninguno. Distinguirlos exige evidencia externa sobre la institución.

Si la proporción superviviente es representativa. No hay garantía general para suponer que lo sea, y en los casos aquí tratados no lo es; en qué medida y en qué dirección no puede determinarse desde los supervivientes, y tampoco puede determinarse desde ellos la afirmación universal de que la supervivencia nunca es representativa.

Las fuentes con identidad bibliográfica no resuelta se excluyeron en vez de conservarse con una salvedad. El argumento restante utiliza únicamente fuentes cuya identidad y función pueden declararse explícitamente.

En qué punto estamos

La evidencia se detiene en cuatro puntos. Las omisiones no se miden desde los fondos supervivientes, porque la evidencia que las mediría es justo lo ausente. Destrucción y abandono, coste y censura, se distinguen por evidencia externa y no por su rastro. Que los supervivientes sean representativos no se resuelve desde ellos, ni tampoco su negación universal. Tres entradas se retiraron en vez de conservarse con una salvedad.

9. Conclusión

La memoria colectiva no es un depósito que se llena. Es un conjunto de instituciones que deciden continuamente qué llevar adelante, y las decisiones se vuelven invisibles al llevarlas.

Un archivo es por tanto dos documentos a la vez: el registro de lo que contiene y el registro no escrito de lo que sus custodios consideraron digno de contener. Leerlo bien significa sostener ambos a la vista y aceptar que el segundo no se lee directamente. No es un defecto que un archivo mejor repare. Es lo que un archivo es, y la respuesta práctica no es fiarse menos de los supervivientes sino saber qué clase de cosa son: la salida de un filtro que no se describe a sí mismo.

En qué punto estamos

Un conjunto de instituciones decide continuamente qué llevar adelante, y esas decisiones se vuelven invisibles en el resultado. Lo que un archivo contiene y lo que sus custodios consideraron digno de contener son dos registros distintos, y solo el primero se lee directamente. Tomar eso por un defecto que reparar confunde lo que la institución es. Qué sabemos, y cómo, es lo que sigue.

Lo que sabemos y cómo lo sabemos

El diccionario profesional define la valoración como identificar los materiales ofrecidos a un archivo que tienen valor suficiente para ser ingresados, con el debate de fondo planteado como qué conservar y qué destruir. La teoría archivística ha descrito al archivero pasando de conservador pasivo a valorador activo dentro de paradigmas cambiantes. Trouillot sitúa el silencio en cuatro momentos: creación del hecho, ensamblaje del hecho, recuperación del hecho y significación retrospectiva. Se ha propuesto leer los archivos como lugares de producción de conocimiento y monumentos de los estados. Existe una norma internacional general de descripción archivística en segunda edición de 2000. Y un estudio

sobre más de un millón de referencias procedentes de más de tres millones y medio de artículos publicados entre 1997 y 2012 informa de que uno de cada cinco artículos de ciencia, tecnología y medicina sufre corrosión de referencias.

Lo que todavía no sabemos

El tamaño y la forma de lo que falta, porque la evidencia que lo mediría es lo ausente. Si una omisión fue deliberada, ya que destrucción, abandono, coste y censura pueden dejar el mismo rastro y lo que los separa es evidencia externa sobre la institución. Si la proporción superviviente es representativa, que no se resuelve desde los supervivientes, y tampoco la afirmación universal de que nunca lo es. Si la explicación de Bagnall sobre Alejandría es la del campo o solo suya: el artículo la da como suya. Las operaciones excluidas no se utilizan para sostener ninguna afirmación de este artículo.

Qué podría hacernos cambiar de opinión

Evidencia externa sobre una institución separaría la destrucción de la negligencia, el coste y la censura, cosa que el archivo por sí solo no puede. Las decisiones de valoración documentadas, en calendarios de eliminación e informes de valoración, recuperan que hubo decisión aunque el contenido destruido sea irrecuperable. Leer directamente los registros oficiales de ISO y CCSDS sacaría esas dos entradas de su estado heredado. Y resolver el desacuerdo sobre la afirmación que sostiene Bagnall zanjaría a cuál corresponde.

Si solo recuerdas una cosa

Los fondos supervivientes de un archivo no revelan todas sus omisiones: lo que sobrevive es evidencia sobre el guardar tanto como sobre lo guardado.

Artículos relacionados

- KNO-008 · *Instrumentos que amplían los sentidos*
- KNO-010 · *La clasificación como herramienta de conocimiento*

Referencias

Bagnall, Roger S. 2002. «Alexandria: Library of Dreams». *Proceedings of the American Philosophical Society* 146 (4): 348–362. URL estable: <https://archive.nyu.edu/bitstream/2451/28263/2/D172-Alexandria%20Library%20of%20Dreams.pdf> CCSDS. 2024. *Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-M-3, Issue 3. Consultative Committee for Space Data Systems. URL estable: <https://ccsds.org/Pubs/650x0m3.pdf> Cook, Terry. 2013. «Evidence, memory, identity, and community: four shifting archival paradigms». *Archival Science* 13 (2): 95–120. DOI 10.1007/s10502-012-9180-7. ICA. 2000. *ISAD(G): General International Standard Archival Description*. 2.ª edición. Ottawa: International Council on Archives, Committee on Descriptive Standards. ISBN 0-9696035-5-X. URL estable: https://www.ica.org/app/uploads/2024/01/CBPS_2000_Guidelines_ISADG_Second-edition_EN.pdf ISO. 2025. *Space Data System Practices: Reference Model for an Open Archival Information System (OAIS)*. ISO 14721:2025, 3.ª edición. Ginebra: International Organization for Standardization. URL estable: <https://www.iso.org/standard/87471.html> Klein, Martin, Herbert Van de Sompel, Robert Sanderson, Harihar Shankar, Lyudmila

Balakireva, Ke Zhou, y Richard Tobin. 2014. «Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot». *PLOS ONE* 9 (12): e115253. DOI 10.1371/journal.pone.0115253. SAA. *Dictionary of Archives Terminology*, entrada «Appraisal». Society of American Archivists. URL estable: <https://dictionary.archivists.org/entry/appraisal.html> Stoler, Ann Laura. 2002. «Colonial Archives and the Arts of Governance». *Archival Science* 2 (1): 87–109. DOI 10.1007/BF02435632. Trouillot, Michel-Rolph. 1995. *Silencing the Past: Power and the Production of History*. Boston: Beacon Press. ISBN 0-8070-4310-9.

KNO-010 · La clasificación como herramienta de conocimiento

KNO · Formas de conocer

Juan Seguro

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

Una clasificación es un instrumento de trabajo, no un retrato del mundo. Se construye para que algo posterior sea posible: reencontrar un ejemplar, contar una defunción, colocar un libro, administrar una población. Este artículo lee la clasificación como una herramienta con fines declarados y costes medibles. Ningún sistema que Bowker y Star examinaron cumple las tres propiedades que tendría uno ideal, la CIE muestra cómo trata un esquema lo que sus categorías no recogen, y agrupar, nombrar e identificar son operaciones que un esquema funde. El caso de la especie muestra una categoría imprescindible y sin resolver a la vez; el rango muestra una convención no comparable entre grupos. La clasificación popular se sitúa junto a la taxonomía científica y no debajo. Aplicadas a personas, adquieren consecuencias que un ejemplar no sufre. La revisión se paga dos veces, y parte de la factura recae sobre todo lo ya contado.

En 60 segundos

La clasificación parece descripción y no lo es. Un esquema es un conjunto de cajas construido para que funcione alguna operación posterior, y sus categorías fueron modeladas por esa operación. Bowker y Star enuncian tres propiedades que tendría una clasificación ideal e informan de que ninguno de los sistemas en funcionamiento que examinaron satisface las tres. Un esquema que no es completo tiene que colocar lo que sus categorías no previeron, y dónde se sitúa ese residuo dice cómo lo trata. Dentro de una sola palabra se esconden tres operaciones: agrupar, nombrar e identificar, que fallan de modo distinto y se reparan de modo distinto. La biología ofrece el caso más agudo, una categoría de especie que se usa en todas partes y que ninguna regla acordada define, y un sistema de rangos cuyos niveles no son comparables entre escarabajos y primates. Los esquemas populares no son borradores previos de los científicos; responden a otra pregunta. Cuando lo clasificado son personas, las propiedades ordinarias adquieren consecuencias, y dos casos documentados muestran un esquema bibliotecario transportando los supuestos de su época y un esquema estatal decidiendo qué era una persona. La revisión cuesta dinero ahora y comparabilidad después.

Por qué importa esta pregunta

Las clasificaciones no se consultan, se habitan. Un código diagnóstico decide como qué se cuenta una defunción, una signatura decide qué encuentra un lector, una categoría de población decide qué le hace un Estado a una persona. Como las categorías fueron modeladas por la operación que tenían que sostener, un esquema leído como informe sobre cómo están divididas las cosas presta al mundo una estructura que vino del trabajo. Ese error de lectura es barato de cometer y caro de mantener: convierte una convención revisable en un hecho y esconde la decisión que una categoría registra.

1. Qué se le pide a una clasificación

Una clasificación no es una descripción. Es un dispositivo para colocar cosas en grupos de modo que sea posible un trabajo posterior, y sus grupos están hechos por el trabajo que tendrán que sostener. Bowker y Star definen una clasificación como una segmentación del mundo y un sistema de clasificación como un conjunto de cajas, metafóricas o literales, en las que pueden ponerse cosas para después hacer algún tipo de trabajo [1]. Lo que conviene retener es el final de esa definición. Un esquema que no sostiene ninguna operación posterior no es una clasificación deficiente. No es una clasificación.

Esto convierte de inmediato la clasificación en un problema de ingeniería con una superficie epistémica. El ingeniero pregunta si el esquema permite que la operación funcione. Quien lee el esquema pregunta qué aspecto exigió esa operación que tuviera el mundo.

Bowker y Star enuncian tres propiedades que tendría una clasificación ideal: opera con principios clasificatorios consistentes y únicos, sus categorías son mutuamente excluyentes y el sistema es completo [2]. Cada una es razonable por separado. Bowker y Star informan de que ninguno de los sistemas en funcionamiento que examinaron satisface las tres, y lo extienden más allá de su muestra como duda argumentada y no como demostración: ningún sistema de clasificación real en funcionamiento de los que examinaron satisface esos requisitos «simples», y dudan de que alguno pudiera hacerlo [3].

Esto es lo primero que hay que entender sobre la clasificación y no es una queja. Los tres requisitos son un patrón con el que leer los esquemas reales, no una especificación que los sistemas examinados cumplieran. Un esquema que los incumple es normal. Un esquema presentado como si los cumpliera está mal descrito.

En qué punto estamos

Una clasificación es un dispositivo construido para que funcione una operación posterior, así que sus grupos fueron modelados por esa operación y no encontrados en el mundo. Tres propiedades harían ideal a una: principios consistentes, categorías mutuamente excluyentes, sistema completo. Ninguno de los sistemas que Bowker y Star examinaron satisface las tres, y dudan de que alguno pudiera. Incumplir ese patrón es lo corriente; presumir de cumplirlo es lo sospechoso.

2. El residuo

La incompletitud es el requisito cuyo incumplimiento se ve con más facilidad, porque un esquema que no es completo tiene que hacer algo con lo que sus categorías no previeron. La CIE muestra una manera explícita en que una clasificación en funcionamiento trata los casos que sus categorías principales no recogen.

En la Clasificación Internacional de Enfermedades el residuo es explícito y está repartido. Bowker y Star observan que la fórmula *not elsewhere classified*, no clasificado en otra parte, aparece a lo largo de toda la CIE pero en ningún caso como categoría de primer nivel [4]. Ese reparto es informativo. Una única caja de primer nivel para lo no clasificado reuniría en un solo sitio lo no nombrado. Repartir el residuo por las ramas mantiene en cambio los efectos de la incertidumbre locales a la rama en que surgen. Eso es lo que la observación establece; lo que un lector concluya del reparto sobre el conjunto del esquema no queda establecido por ella.

Una categoría residual no es un hueco de la clasificación. Es una pieza en funcionamiento de ella, y leer bien un esquema consiste en leer dónde están sus residuos y cuánto tráfico soportan.

En qué punto estamos

La incompletitud es el requisito cuyo incumplimiento se ve con más claridad, porque lo que las categorías no previeron hay que colocarlo igual. En la CIE la fórmula residual recorre las ramas y nunca se sitúa en el primer nivel, lo que mantiene la incertidumbre local donde surge. Ese reparto es lo que la observación establece. Leer bien un esquema es ver dónde están sus residuos y cuánto tráfico soportan.

3. Tres operaciones que un solo esquema oculta

Clasificación se usa con regularidad como una sola palabra para tres operaciones distintas, y separarlas elimina buena parte de la confusión.

Agrupar decide qué cosas van juntas. Es la operación que describen las secciones anteriores.

Nombrar adjudica una etiqueta a un grupo. Se rige por convenciones propias, y esas convenciones pueden cambiar sin que cambie una sola agrupación. Nada en el mundo decide si un grupo se llama de una manera o de otra; lo decide un código de nomenclatura.

Identificar decide a qué grupo pertenece una cosa concreta. Funciona con la evidencia disponible en el momento de la decisión, que suele ser mucho menor que la que estableció el grupo.

Las tres se separan en la práctica. Dos personas pueden coincidir en los grupos, discrepar en los nombres y tener ambas razón respecto al código que siguen. Dos personas pueden coincidir en los nombres y situar el mismo ejemplar de forma distinta porque la clave de identificación resuelve lo que pueden ver y no aquello por lo que el grupo está definido. **Cuando se critica un esquema conviene preguntar cuál de las tres operaciones se está criticando en realidad**, porque los remedios no tienen relación entre sí.

En qué punto estamos

Una sola palabra cubre tres operaciones. Agrupar decide qué va junto, nombrar adjudica una etiqueta bajo un código que puede cambiar sin mover nada, e identificar sitúa una cosa concreta con la evidencia que haya a mano. Se separan en la práctica: el acuerdo sobre los grupos sobrevive al desacuerdo sobre los nombres. Clasificar una crítica en las tres compensa, porque sus remedios no tienen relación.

4. El problema de la especie

La biología ofrece el caso más claro de una clasificación cuya categoría más usada carece de definición acordada. Wells y colaboradores afirman que las especies siguen siendo controvertidas en cuanto a definición, delimitación y realidad [5]. No es una posición minoritaria ni reciente. Las especies se cuentan, se protegen por ley, se declaran extintas y se usan como unidad de una parte muy grande de la investigación biológica, y la unidad misma está en discusión.

Su propuesta merece enunciarse con precisión porque no es una retirada. Sostienen que tratar la especie como una heurística puede ser compatible con una definición teórica de qué son las especies [6], siendo el contenido teórico que las especies son conjuntos de individuos estrechamente

emparentados que responden de manera similar a conjuntos comparables de fuerzas evolutivas y ecológicas [6]. Una heurística en este sentido no es una confesión de fracaso. Es una afirmación sobre qué tipo de instrumento es la categoría: uno que hace un trabajo fiable en un rango amplio de casos sin ser reducible a una sola regla de decisión.

El caso de la especie muestra algo general. Una categoría puede ser imprescindible y estar sin resolver a la vez, y lo imprescindible no es prueba de que se haya encontrado la definición.

En qué punto estamos

La especie es una unidad de gran parte de la biología y está en discusión, mientras esa misma unidad se cuenta, se protege por ley y se declara extinta. Wells y colaboradores la llaman controvertida en definición, delimitación y realidad, y sostienen que tratar la categoría como heurística puede convivir con una explicación teórica de qué son las especies. Necesitarla sin cesar no prueba que se haya encontrado su definición.

5. El rango es una convención

Por encima de la especie la dificultad cambia de forma. La pregunta ya no es qué es la categoría, sino qué significa su nivel.

Lee y Skinner afirman que no hay criterios generalmente aceptados para especificar el rango al que debe referirse un taxón, y que los criterios propuestos presentan dificultades; añaden que los taxones asignados a un mismo rango rara vez son entidades equivalentes entre grupos distintos [7]. Una familia de escarabajos y una familia de primates no son cantidades comparables. Comparten una palabra.

La nomenclatura filogenética respondió proponiendo nombrar clados sin rangos. Benton informa de que la mayoría de sus partidarios sostiene que los rangos deberían abandonarse, señalando a la vez que el propio PhyloCode no lo exige, y objeta que los rangos aportan información utilizable sobre la inclusividad relativa de los clados y sirven de aproximación práctica en trabajos de macroevolución y biodiversidad: por qué abandonar los rangos, pregunta, cuando al hacerlo no se gana nada salvo pureza semántica [8].

Ninguna de las dos posiciones discute que los rangos no sean comparables entre grupos. Lo que informan Lee y Skinner y lo que argumenta Benton difieren sobre si una convención que se sabe no comparable merece conservarse por el trabajo que hace. Son dos posiciones enunciadas entre otras, no las dos mitades del campo. Esa es una pregunta sobre la herramienta y no sobre la naturaleza, y es la forma correcta de pregunta para cualquier clasificación.

En qué punto estamos

Por encima de la especie la pregunta pasa de qué es la categoría a qué significa su nivel. Lee y Skinner afirman que no hay criterios generalmente aceptados para asignar rango y que los taxones de un mismo rango rara vez son equivalentes entre grupos. Benton no lo discute y sostiene que la convención sigue valiendo la pena. El desacuerdo es sobre la herramienta, la pregunta pertinente ante cualquier esquema.

6. La clasificación popular, al lado y no debajo

La taxonomía científica es una clasificación de los seres vivos entre muchas, y las demás no son borradores previos de ella.

El trabajo etnobiológico ha encontrado estructura recurrente. Atran informa de que Berlin estableció una terminología estándar de rangos etnobiológicos: el rango de reino popular, el rango de forma de vida, el rango genérico o de especie genérica, el rango específico popular y el rango varietal popular [9]. La recurrencia es el hallazgo. Dice que grupos humanos que trabajan de forma independiente llegan a jerarquías de profundidad parecida, y que un nivel en particular concentra la mayor parte del nombrar y del aprender.

El ajuste con los taxones científicos es parcial y es desigual de una manera concreta. Atran afirma que las especies genéricas se corresponden a menudo con géneros o especies científicos, al menos para los organismos que los humanos perciben con más facilidad, como los vertebrados grandes y las plantas con flor, y que en ocasiones se corresponden en cambio con fragmentos locales de familias, órdenes o, sobre todo en invertebrados, taxones aún superiores [10]. Atran lee esa desigualdad como seguimiento de la saliencia perceptiva y no de la filogenia, y esa lectura es su aportación a un debate y no un resultado zanjado. Se recoge aquí como el patrón que él informa en los casos que examinó, no como una regla que fije de antemano dónde convergerán dos esquemas cualesquiera.

Ni la recurrencia ni el desajuste sostienen una escalera evolutiva. Una clasificación que agrupa por lo que un organismo hace, por dónde se encuentra o por qué puede hacerse con él responde a una pregunta distinta de la que responde una que agrupa por descendencia, y responder a otra pregunta no es responder mal a la misma.

En qué punto estamos

Grupos humanos que trabajan por separado llegan a jerarquías de profundidad parecida, y Atran expone la terminología estándar que Berlin estableció para sus rangos. La correspondencia con los taxones científicos es parcial y desigual, y la lectura que Atran hace de ella es su aportación a un debate, no un resultado zanjado. Un esquema que agrupa por uso o hábitat responde a otra pregunta que uno que agrupa por descendencia.

7. Cuando lo clasificado es una persona

Todo lo anterior se aplica a ejemplares. Cuando los objetos son personas, los mismos rasgos estructurales adquieren consecuencias que un ejemplar no sufre.

El término de Bowker y Star para el mecanismo central es *torque*, torsión. Donde una vida individual y una clasificación institucional no encajan, los autores describen una presión que corre en los dos sentidos: los autores llaman torsión a ese aglomerado, un retorcimiento de líneas temporales que tiran unas de otras y doblan o tuercen tanto la biografía del paciente como el proceso de metrificación [11]. Tanto la biografía como el proceso de medición quedan torcidos por el encuentro, en la formulación de los propios autores, y lo que se registra después lo arrastra.

Dos casos documentados muestran el alcance.

Clasificación bibliotecaria. Thornton rastrea cómo las teorías sobre la clasificación humana procedentes de la Antigüedad informaron las ideas de raza y etnia de los siglos XVIII y XIX, cómo esas ideas quedaron incrustadas en la Clasificación Decimal Dewey y cómo el tratamiento de esas materias

cambió a lo largo de cuatro ediciones y 130 años, reflejando la presentación de visiones sociales dominantes cambiantes [12]. El esquema no se limitó a contener los supuestos de su época. Los transportó hacia adelante, edición tras edición, hasta la ordenación de colecciones construidas mucho después de que aquellos supuestos se abandonaran.

Clasificación estatal de la población. El estudio de Posel sobre la Sudáfrica del siglo XX describe criterios de clasificación que combinaban la apariencia y la aceptación general con hábitos, educación, habla y comportamiento, y constata que los funcionarios disponían de una latitud considerable para interpretar los criterios burocráticos de clasificación racial de modo que podían recurrir de forma amplia e idiosincrásica a repertorios populares de estereotipos biológicos [13]. Lo que importa aquí es estructural y es grave. Una clasificación que no puede aplicarse de forma consistente no deja por ello de operar. Opera a través de la discrecionalidad de quien la aplica, y el registro resultante muestra una categoría donde en realidad hubo una decisión.

Ninguno es un uso indebido de un instrumento neutral, y los dos no son el mismo caso. Un esquema bibliotecario arrastró los supuestos de su época hasta la disposición de las colecciones; un esquema estatal decidió qué era una persona a efectos administrativos. Ambos muestran las propiedades ordinarias descritas en las secciones 1 a 3 operando sobre objetos a los que ser clasificados puede dañar, y difieren en qué hizo la clasificación.

En qué punto estamos

Con personas estas propiedades adquieren consecuencias que un ejemplar nunca sufre. Bowker y Star llaman torsión a una presión de doble sentido que dobla a la vez la biografía y el proceso de medición. Un esquema bibliotecario llevó los supuestos de su época a colecciones muy posteriores; uno estatal decidió qué era una persona. Ninguno es un instrumento neutral del que se abuse, y lo que hizo cada uno difiere.

8. La revisión y quién la paga

Una clasificación usada durante suficiente tiempo tiene que revisarse. La revisión se paga dos veces: en el presente, mediante la implantación y el reciclaje de quienes codifican, y de forma retrospectiva, sobre los datos ya recogidos.

La Clasificación Internacional de Enfermedades es el instrumento más claro para verlo porque se revisa de forma deliberada, a intervalos largos y por un organismo internacional. Su undécima revisión fue adoptada por la 72.^a Asamblea Mundial de la Salud el 28 de mayo de 2019, en una resolución que adopta la undécima revisión de la Clasificación Estadística Internacional de Enfermedades y Problemas de Salud Conexos, CIE-11, para que entre en vigor el 1 de enero de 2022, con sujeción a disposiciones transitorias [14].

El coste de un cambio así es medible. Janssen y Kunst examinaron la mortalidad por causas en seis países europeos entre 1950 y 1999 y hallaron que, pese a usar una tabla de concordancia basada en códigos CIE de tres dígitos, en el 10,8 % de las revisiones de códigos de mortalidad por causas persistieron discontinuidades significativas atribuidas a las propias revisiones de la CIE, y concluyeron que las discontinuidades derivadas de los cambios de codificación pueden producir cambios sustanciales en las tendencias a largo plazo de la mortalidad por causas [15].

Leído con cuidado, eso es una afirmación sobre los límites de la traducción entre esquemas. En las revisiones que esos seis países atravesaron entre 1950 y 1999, se aplicó una tabla de concordancia como la mejor reparación disponible y aproximadamente una revisión de cada nueve dejó aún una

discontinuidad que los autores atribuyeron a la revisión misma y no a ningún cambio en la salud. La cifra corresponde a esas transiciones y no es una tasa para las revisiones en general. **Una serie temporal larga construida a través de revisiones de una clasificación contiene por tanto dos señales superpuestas, una sobre el mundo y otra sobre el esquema**, y solo la primera suele ser aquella por la que la serie se lee.

En qué punto estamos

La revisión se paga dos veces: ahora, mediante implantación y reciclaje, y hacia atrás, sobre lo ya registrado. Janssen y Kunst midieron la segunda factura en seis países a lo largo de cinco décadas: la concordancia reparó lo que pudo y cerca de una revisión de cada nueve seguía arrastrando una discontinuidad atribuida al cambio de codificación y no a la salud. La cifra es suya, no una tasa general.

9. Leer bien una clasificación

Lo que sostienen las secciones anteriores no es escepticismo sobre la clasificación. Es un conjunto breve de preguntas que plantear a cualquier esquema en uso.

¿Para qué operación se construyó? Las categorías fueron modeladas por esa operación y encajarán peor en otras.

¿Dónde está el residuo y cuánto pasa por él? Una categoría residual grande o creciente es el esquema informando de su propia tensión.

¿Cuál está en juego: agrupar, nombrar o identificar? Las tres fallan de modo distinto y se reparan de modo distinto.

¿Está resuelta la categoría? Imprescindible y definida son propiedades distintas, como muestra el caso de la especie.

¿Significan algo los niveles entre ramas? Con frecuencia no, y la palabra compartida lo oculta.

¿Quién o qué se está clasificando, y qué le ocurre a lo que no encaja? Con ejemplares la respuesta es una categoría residual. Con personas puede ser torsión.

¿Cuánto costó la última revisión y a quién? La factura la paga todo lo ya registrado.

Una clasificación es una buena herramienta en el sentido corriente: hace bien un trabajo determinado, deforma aquello para lo que no fue construida y hay que mantenerla. Tratarla como un informe sobre cómo está dividido el mundo atribuye al mundo una estructura que vino del trabajo.

En qué punto estamos

El artículo termina donde un lector puede usarlo: una lista breve de preguntas para cualquier esquema en uso, sobre la operación para la que se construyó, su residuo, cuál de las tres operaciones está en juego, si la categoría está resuelta, si los niveles significan algo entre ramas, qué le ocurre a lo que no encaja y cuánto costó la última revisión. Qué sabemos, y cómo, es lo que sigue.

Lo que sabemos y cómo lo sabemos

Bowker y Star definen un sistema de clasificación como un conjunto de cajas en las que poner cosas para después hacer algún tipo de trabajo, enuncian tres propiedades que tendría una clasificación ideal e informan de que ninguno de los sistemas en funcionamiento que examinaron satisface las tres. Observan que la fórmula residual recorre toda la CIE pero nunca como categoría de primer nivel, y llaman torsión al retorcimiento mutuo de una biografía y un proceso de metrificación. Wells y colaboradores informan de que la especie sigue siendo controvertida en definición, delimitación y realidad, y proponen tratarla como heurística. Lee y Skinner no encuentran criterios generalmente aceptados de rango; Benton sostiene que los rangos siguen haciendo un trabajo utilizable. Atran expone los rangos populares de Berlin y su correspondencia parcial con los taxones científicos. Thornton rastrea el tratamiento de Dewey a lo largo de cuatro ediciones y 130 años, Posel documenta la latitud de los funcionarios en la clasificación racial sudafricana, la 72.^a Asamblea Mundial de la Salud adoptó la CIE-11 en 2019, y Janssen y Kunst hallaron discontinuidades persistentes en el 10,8 % de las revisiones de códigos en seis países.

Lo que todavía no sabemos

Qué es una especie, cuando la categoría se usa en todas partes y su definición está en discusión, y la propuesta heurística es una manera de trabajar y no una resolución. Si conviene conservar los rangos, que Lee y Skinner y Benton responden de modo distinto coincidiendo en que los rangos no son comparables entre grupos. Hasta dónde se corresponden las especies genéricas populares y los taxones científicos, que Atran lee como saliencia perceptiva y ofrece como aportación a un debate. Y cuánto de una serie larga de mortalidad es el esquema y no el mundo, ya que el 10,8 % corresponde a las transiciones estudiadas y no es una tasa para las revisiones en general.

Qué podría hacernos cambiar de opinión

Un criterio operativo de rango que se sostuviera entre grupos convertiría la defensa práctica de Benton en una más fuerte, y su ausencia continuada es lo que informan Lee y Skinner. Trabajo etnobiológico sobre grupos cuyos entornos difieran mucho de los ya estudiados pondría a prueba si la recurrencia que describe Atran es tan general como parece. Un trabajo de concordancia que separase la discontinuidad de codificación del cambio real, revisión por revisión y no en agregado, diría cuánto de una serie larga es el esquema. Y una clasificación que documentara el tráfico de sus propios residuos dejaría ver la tensión en lugar de deducirla.

Si solo recuerdas una cosa

Una clasificación se construye para hacer posible una operación, así que sus categorías no pueden leerse como un retrato neutral del mundo.

Artículos relacionados

- KNO-009 · *Archivos, bibliotecas y memoria colectiva*

Referencias

Los números entre corchetes del cuerpo son números de afirmación y no de referencia. Cada uno remite a una fila del registro de afirmaciones, que nombra la fuente de abajo que la sostiene.

Atran, Scott. 1998. «Folk biology and the anthropology of science: Cognitive universals and cultural particulars». *Behavioral and Brain Sciences* 21 (4): 547–569. URL estable: <https://philpapers.org/rec/ATRFAT>

Benton, Michael J. 2007. «The Phylocode: Beating a dead horse?». *Acta Palaeontologica Polonica* 52 (3): 651–655. URL estable: <https://www.app.pan.pl/archive/published/app52/app52-651.pdf>

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting things out: classification and its consequences*. Cambridge, Mass.: MIT Press. ISBN 0-262-02461-6.

Janssen, Fanny, and Anton E. Kunst. 2004. «ICD coding changes and discontinuities in trends in cause-specific mortality in six European countries, 1950-99». *Bulletin of the World Health Organization* 82 (12): 904–913. **Sin DOI localizado**. URL estable: <https://scielosp.org/article/bwho/2004.v82n12/904-913/>

Lee, Michael S.Y., and Adam Skinner. 2007. «Stability, ranks, and the PhyloCode». *Acta Palaeontologica Polonica* 52 (3): 643–650. URL estable: <https://www.app.pan.pl/archive/published/app52/app52-643.pdf>

Posel, Deborah. 2001. «Race as Common Sense: Racial Classification in Twentieth-Century South Africa». *African Studies Review* 44 (2): 87–113. DOI 10.2307/525576.

Thornton, Lisa. 2024. «How Did We Get Here? Race and Ethnicity in Dewey Decimal Classification». *Knowledge Organization* 51 (6): 414–424. DOI 10.5771/0943-7444-2024-6-414.

Wells, Tom, Tom Carruthers, Pablo Muñoz-Rodríguez, Alex Sumadijaya, John R I Wood, and Robert W Scotland. 2022. «Species as a Heuristic: Reconciling Theory and Practice». *Systematic Biology* 71 (5): 1233–1243. DOI 10.1093/sysbio/syab087.

World Health Organization. 2019. *Eleventh revision of the International Classification of Diseases*. Resolución WHA72.15, punto 12.7 del orden del día, 72.^a Asamblea Mundial de la Salud, 28 de mayo de 2019. URL estable: https://cdn.who.int/media/docs/default-source/classification/icd/icd11/a72_29-en_icd-11-adoption.pdf

KNO-011 · Medir antes de teorizar

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

Suele decirse que la teoría indica a los científicos qué medir, y con frecuencia lo hace. Este artículo sigue la dirección contraria, en la que los números llegan primero y se sostienen hasta que una explicación los alcanza, y no trata ninguna como la regla. Abre con la pregunta que KNO-010 entregó, si clasificar en categorías cuenta ya como medir, y nombra las dos respuestas sin adoptar ninguna. Toma después tres casos en los que el número llegó antes: un termómetro en el que había que confiar antes de poder comprobarlo, medidas en longitudes de onda largas que una fórmula célebre se escribió para ajustar, y un residuo orbital que ninguna teoría disponible explicaba. Cierra con el movimiento inverso de 2019, cuando una constante medida un siglo pasó a ser una definición. El diseño experimental, el análisis de la incertidumbre y la estructura de los registros corresponden a artículos posteriores.

En 60 segundos

Los dos órdenes existen y ninguno es el que este campo sigue: a veces la teoría nombra una magnitud y un experimento va a buscarla, y a veces los números se acumulan hasta que una teoría tiene que reproducirlos. Si clasificar cuenta ya como medir depende de qué definición esté en vigor, y aquí se nombran la permisiva y la restrictiva sin adoptar ninguna. Hay que confiar en un instrumento antes de que exista nada capaz de comprobarlo, y en los episodios que se tratan el círculo no se evitó en lugar de romperse. Una fórmula puede escribirse para acertar un blanco que la medida ya había puesto: una comunicación sobre el comportamiento en longitudes de onda largas llegó antes y la ecuación después, lo que es un orden de los hechos y no una explicación de qué causó qué. Un residuo informa del tamaño del fallo de una teoría y obliga solo a las sucesoras que cubran el mismo dominio observacional, con la precisión que allí importa. En 2019 una constante medida pasó a ser definida, y el valor definitorio no tiene incertidumbre alguna; lo que la carga ahora es la realización de la unidad.

Por qué importa esta pregunta

Un número ofrecido como medida no se explica solo. Leído sin preguntar qué validó el instrumento, si la magnitud se teorizó antes, y si la cifra es un valor o un residuo, ese mismo número acabará sosteniendo conclusiones que no puede cargar. Las distinciones que aquí se hacen son las que permiten a quien lee ver que una fórmula respondía ante los datos antes de significar nada, que una corrección pequeña y estable puede sobrevivir a todos los intentos de explicarla, y que una constante puede dejar de medirse por decisión y no por hallazgo.

1. Dos órdenes, y por qué ninguno es la regla

Existe un relato ordenado en el que la teoría va primero. Una teoría afirma que cierta magnitud existe, dice aproximadamente cuánto debería valer, y a continuación se diseña un experimento para encontrarla. Buena parte de la física moderna funciona así, y el relato no es falso.

Existe un relato igual de ordenado en sentido contrario. Se mide algo porque se puede, los números se acumulan, aparece una pauta en ellos, y acaba construyéndose una teoría que tiene que reproducir esa pauta o ser descartada. Buena parte de la historia de la física y de la química funciona también así, y ese relato tampoco es falso.

Lo falso es tratar cualquiera de los dos como el orden de la ciencia. Las dos direcciones coexisten, se alternan dentro de un mismo programa de investigación y a veces operan sobre la misma magnitud en momentos distintos. Este artículo desarrolla la segunda dirección porque es la que suele quedarse fuera, no porque sea la verdadera, y la sección 6 muestra la primera imponiéndose de forma concluyente en un caso donde la segunda había regido durante un siglo.

En qué punto estamos

Dos relatos ordenados van en sentidos opuestos y los dos describen práctica real, y por eso ninguno es el orden que este campo sigue. Que la teoría nombre una magnitud y un experimento vaya a por ella ocurre; que los números se acumulen hasta que una teoría tenga que reproducirlos ocurre también, a veces sobre la misma magnitud. Aquí se desarrolla el segundo porque suele quedar fuera.

2. Qué cuenta como medir, y la pregunta que KNO-010 dejó abierta

KNO-010 se detuvo en la categoría y entregó una pregunta que no argumentó: si asignar una cosa a una categoría es ya una medida. Hay que responderla antes que nada aquí, porque de la respuesta depende qué cuenta como caso.

La respuesta permisiva es la más citada. Stevens define la medida, en el sentido más amplio, como la asignación de numerales a objetos o sucesos de acuerdo con reglas [1]. Sobre esa definición sus cuatro tipos de escala se siguen de lo que la asignación permite hacer. Su tabla los separa por la operación empírica básica de cada uno: la escala nominal por determinación de igualdad, la ordinal por determinación de mayor o menor, la de intervalo por determinación de igualdad de intervalos o diferencias, y la de razón por determinación de igualdad de razones [2].

La respuesta restrictiva objeta que así la pregunta deja de poder plantearse. Michell observa que, bajo la posición de Stevens, cualquier regla para hacer asignaciones numéricas define siempre al menos una escala nominal [3]. Si cualquier regla produce una medida, entonces nunca se somete a prueba si el atributo estudiado es cuantitativo de entrada, y ese es exactamente el cargo de Michell contra un siglo de psicología cuantitativa: la cuestión científica de la aditividad de los atributos implicados casi nunca se abordó [4].

Ambas posiciones quedan registradas y no se adopta ninguna. La consecuencia para este artículo es solo procedimental, y basta con ella. **Bajo la definición permisiva, clasificar es medir en la escala más débil; bajo la restrictiva, no es medir en absoluto mientras no se muestre que el atributo es cuantitativo.** Todo lo que sigue trata de magnitudes que satisfacen también la definición restrictiva, de modo que los casos de abajo se sostienen bajo cualquiera de las dos lecturas. Cuando un artículo necesite que la distinción haga trabajo real, debe decir qué definición está usando, porque las dos discrepan sobre qué se ha logrado cuando se produce un número.

En qué punto estamos

Si clasificar cuenta ya como medir depende de la definición en vigor. La permisiva admite cualquier asignación de numerales regida por reglas y ordena los resultados por lo que la asignación permite; la restrictiva objeta que así nunca se somete a prueba si el atributo es cuantitativo. Las dos quedan registradas y no se adopta ninguna, y todo lo de abajo satisface la lectura más estricta.

3. Validar un instrumento sin un patrón detrás

La primera dificultad de medir antes de teorizar no es la teoría. Es que hay que confiar en el instrumento antes de que exista nada capaz de comprobarlo.

El termómetro es el ejemplo permanente. Para saber si un termómetro marca correctamente hace falta una temperatura conocida, y para conocer una temperatura hace falta un termómetro. El trabajo de los siglos XVIII y XIX sobre los puntos fijos fue un intento de romper ese círculo desde dentro, y es el asunto del estudio de Chang, que su editorial describe como un examen de cómo aquellos científicos lograron evaluar la fiabilidad y la exactitud de esos instrumentos sin apoyarse circularmente en los propios instrumentos [5].

El círculo no es un escándalo, y en los casos que este artículo trata no se evitó. Varios patrones de medida fueron establecidos en algún momento por instrumentos que después ese mismo patrón sirvió para juzgar. Si eso vale para todo patrón no se establece aquí y no se afirma. Lo que puede exigirse a un proceso así no es que parta de la certeza, sino que cada vuelta deje a la siguiente en mejor posición, y que las mejoras sean visibles en lugar de supuestas. Este artículo no desarrolla la epistemología general de ese proceso, que es el objeto del libro de Chang y excede el alcance de la serie. **Lo que aquí se usa de ese libro se limita a lo comprobado directamente sobre su resumen y su índice oficiales.** Su cuerpo no se leyó, de modo que nada de esta sección le atribuye un argumento más allá de la descripción que da su editorial.

Una consecuencia importa aquí. Una magnitud puede medirse bien, de forma consistente y útil durante mucho tiempo antes de que exista teoría alguna de qué se está midiendo. La temperatura se midió durante un periodo largo antes de que la termodinámica diera cuenta de ella, y este artículo no fija la duración de ese periodo, que depende de qué se cuente como medida y de qué exposición de la termodinámica se tome como llegada. Las medidas no esperaban a la teoría; fue la teoría la que tuvo que acomodarlas.

En qué punto estamos

Hay que confiar en un instrumento antes de que exista nada capaz de comprobarlo, y el termómetro es el caso permanente. En los episodios que se tratan el círculo no se evitó; hasta dónde llega eso no se zanja aquí. Lo que un proceso así debe no es certeza al empezar sino mejora visible en cada vuelta. Una magnitud puede medirse bien mucho antes de que algo la explique.

4. Una fórmula construida para ajustar las medidas

El caso más claro de una teoría construida para reproducir un número es también uno de los de mayores consecuencias, y su relato habitual invierte el orden.

Hacia 1900 existían varias leyes de radiación en competencia para la emisión de un cuerpo negro. González de Arrieta registra que todas eran compatibles con las propiedades esperadas de la radiación del cuerpo negro pero que, de forma decisiva, cada una predecía comportamientos distintos

para valores grandes del producto λT [6]. El desacuerdo quedaba por tanto localizado con precisión: las leyes eran indistinguibles en el rango accesible y solo se separaban en longitudes de onda largas, lo que es la descripción de un experimento pendiente de hacer.

Rubens y Kurlbaum lo hicieron. Se propusieron mejorar un experimento de rayos residuales ya existente, usando cuatro cavidades de cuerpo negro que abarcaban un intervalo de temperaturas de -180 a 1600 °C [7]. El resultado fue concluyente primero en sentido negativo: ninguna de aquellas expresiones reproducía con precisión los datos observados para todas las temperaturas [8].

Lo que ocurrió después es el punto de esta sección. El 7 de octubre de 1900 Rubens comunicó a Planck que el límite de la ley de radiación para λT grande era una función lineal de la temperatura, y Planck rehízo las propiedades termodinámicas de su sistema de cuerpo negro de base empírica y produjo una ecuación que predecía los fenómenos observados [9]. **La secuencia es la que registra la fuente, y este artículo la lee como secuencia:** la comunicación llegó antes y la ecuación después, lo cual no es por sí solo una explicación de qué causó qué. Rubens confirmó el ajuste, y los dos resultados se presentaron juntos: Rubens contestó que la ecuación propuesta ajustaba los resultados a la perfección, y tanto Planck como Rubens y Kurlbaum presentaron sus resultados en la misma sesión de la Sociedad de Física el 19 de octubre [10].

La fórmula respondía ante las medidas desde el momento en que se escribió. Su interpretación física, la que la convirtió en un resultado fundacional de la teoría cuántica, llegó después y hubo de conciliarse con una fórmula que ya existía y ya funcionaba. Eso es medir antes de teorizar en su forma más fuerte, y conviene notar que el artículo no afirma que la medida causara la teoría. Afirma algo más débil y más defendible: **la teoría tenía un blanco que acertar, y el blanco lo puso la medida.**

En qué punto estamos

Las leyes de radiación rivales coincidían donde los instrumentos llegaban y se separaban solo en longitudes de onda largas, lo que describe un experimento pendiente. Se hizo, ninguna de las expresiones lo sobrevivió, y la ecuación que sí lo hizo llegó tras una comunicación sobre el comportamiento del límite. Ese orden de los hechos es el que registra la fuente. La fórmula respondía ante los números desde que existió.

5. Un residuo que estableció la medida y no explicó la teoría

Una medida puede además producir un número que ninguna teoría explica, y mantenerlo en pie durante décadas.

En 1859 Le Verrier dirigió a Faye una carta sobre la teoría de Mercurio y el movimiento del perihelio de ese planeta, publicada en los Comptes rendus de la Académie des sciences [11]. Su hallazgo era cuantitativo y estrecho. Tal como está impreso el pasaje, bastó aumentar en treinta y ocho segundos de arco el movimiento secular del perihelio para representar todos los pasos observados con un margen inferior a un segundo [12].

Leída con cuidado, esa frase tiene una estructura inusual para un descubrimiento. No informa de un objeto nuevo ni de un efecto nuevo. **Informa del tamaño de la corrección que hay que aplicar a una teoría existente para que esa teoría se ajuste a las observaciones**, es decir, informa de la magnitud exacta del fallo de la teoría. La corrección era pequeña, era estable y no se iba a marchar.

Un residuo de esta clase es un objeto científico peculiarmente duradero, y conviene enunciar de forma transferible la regla para leerlo. Un residuo es una estimación que depende de las observaciones que hay detrás, de los parámetros usados y del modelo del que es residuo. Restringe por tanto solo a las teorías sucesoras que pretendan cubrir el mismo dominio observacional, y solo con la precisión que allí resulte pertinente. No es una anomalía en el sentido laxo de algo raro; es un número, establecido por medida, que una teoría sucesora que cubra el mismo terreno tiene que reproducir dentro de esa precisión. Sobrevivió a todos los intentos de explicarlo ajustando el inventario de masas del sistema solar, y seguía ahí cuando más de medio siglo después llegó una teoría que lo predecía. **No seguía ahí sin cambios.** Laskar recoge la cifra de Le Verrier como unos treinta y ocho segundos de arco por siglo, obtenida de veintiuna observaciones de tránsitos entre 1697 y 1848, frente a los aproximadamente cuarenta y tres segundos de arco por siglo que da la teoría posterior [12]. Son valores próximos y no son el mismo número, y la diferencia es la que hay entre una estimación y una predicción, no un defecto de ninguna de las dos. Este artículo se detiene deliberadamente en el residuo y no trata la explicación posterior, que pertenece a la historia de la gravitación y no a la epistemología de la medida.

En qué punto estamos

Una medida puede dejar un número que ninguna teoría explica y sostenerlo décadas. No se informó de un objeto nuevo sino del tamaño de la corrección que una teoría existente necesitaba, la magnitud de su fallo. Una cifra así es una estimación, así que obliga solo a las sucesoras que cubran el mismo terreno con la precisión que allí importa, y el valor posterior es próximo sin ser el mismo.

6. Cuando una constante medida pasa a ser definida

Las tres secciones anteriores van todas de la medida hacia la teoría. El cambio grande más reciente en metrología va en sentido contrario, y conviene ponerlo al lado para que el artículo no sea una historia de dirección única.

Hasta 2019 el kilogramo era un objeto. Como afirma el apéndice oficial del folleto del SI dedicado al kilogramo, la definición anterior del kilogramo se basaba en la masa del prototipo internacional del kilogramo inmediatamente después del procedimiento de limpieza prescrito [13]. Todo lo expresado en kilogramos era, en última instancia, una comparación con un artefacto.

La Resolución 1 de la 26.^a Conferencia General de Pesas y Medidas, adoptada en 2018 y en vigor desde el 20 de mayo de 2019, sustituyó aquello por una constante. El kilogramo queda definido tomando el valor numérico fijo de la constante de Planck h como $6,626\,070\,15 \times 10^{-34}$ cuando se expresa en la unidad J s, que es igual a $\text{kg m}^2 \text{s}^{-1}$ [14].

La inversión es completa y conviene enunciarla sin rodeos. La constante de Planck se midió, con precisión creciente, durante un periodo largo antes de que su valor quedara fijado; este artículo no pone cifra a ese periodo. Ya no se mide, porque ya no tiene un valor que encontrar: su valor numérico está ahora fijado por definición, y la declaración oficial es inequívoca al afirmar que los valores numéricos de las siete constantes definitorias no tienen incertidumbre [15]. La incertidumbre no desapareció. Se desplazó. Donde antes había una constante incierta y una unidad exacta, ahora hay una constante exacta y una unidad cuya realización carga con la incertidumbre.

La propia resolución registra qué le ocurrió al prototipo, y conviene leerlo con exactitud. El Apéndice 2 de la propia Resolución 1, bajo *Status of constants previously used in the former definitions*, afirma que desde el 20 de mayo de 2019 la masa del prototipo internacional del kilogramo

es de 1 kg dentro de una incertidumbre típica relativa igual a la del valor recomendado de la constante de Planck en el momento de adoptarse la resolución, $1,0 \times 10^{-8}$, y que a partir de entonces su valor se determina experimentalmente [16]. **Esto no es una cantidad conservada que cambie de manos.** La constante definitoria no tiene incertidumbre alguna, que es el enunciado aparte de más arriba con su propia fuente; lo que este apéndice añade es que el artefacto pasa a tener una masa determinada experimentalmente que carga con una incertidumbre heredada, de modo que se *desplazó*, en el párrafo anterior, describe dónde se hace el trabajo y no una cantidad que se pase de mano en mano.

Esto es una convención metrológica, adoptada deliberadamente, y no un caso de la teoría diciendo a la medida qué hacer. Un comité decidió fijar un número que hasta entonces se buscaba, y fue posible solo porque la medida había hecho la constante lo bastante conocida como para que fijarla no cambiara nada en la práctica. Las dos direcciones no rivalizan aquí. La segunda se la ganó la primera.

En qué punto estamos

El cambio grande más reciente fue en sentido contrario. Una constante medida durante un siglo quedó fijada por decisión, y el valor definitorio no tiene incertidumbre porque ya no hay nada que encontrar. El prototipo conserva una masa determinada experimentalmente que carga con una incertidumbre relativa heredada. Nada se pasó de mano en mano: lo que cambió es dónde se hace el trabajo de determinar.

7. Leer bien una medida

Lo que sostienen las secciones anteriores es un conjunto breve de preguntas que plantear a cualquier número ofrecido como medida.

¿Qué definición de medida está en juego? Bajo la permisiva, una clasificación con numerales adheridos cuenta. Bajo la restrictiva, nada cuenta mientras no se muestre que el atributo es cuantitativo. Las dos discrepan sobre qué se ha establecido.

¿Qué validó el instrumento, y qué validó eso? Toda cadena termina en algún sitio, y lo interesante es si el último eslabón se examinó o se supuso.

¿Se teorizó la magnitud antes de medirla? Si no, las medidas no confirman nada y no deben leerse como si lo hicieran.

¿Este número es un valor o un residuo? Un residuo informa del tamaño del fallo de una teoría, y restringe a las sucesoras de un modo en que una medida corriente no lo hace.

¿Se sigue midiendo la constante? Algunas no, por decisión. Una constante definida no tiene incertidumbre porque la incertidumbre se ha desplazado a otro sitio, no porque se haya eliminado.

¿Qué haría falta para que este número cambiara? Un instrumento mejor, una redefinición o una reinterpretación son respuestas distintas con consecuencias distintas, y un número que no puede cambiar por ninguna de las tres probablemente no es una medida.

Medir antes de teorizar no es una actividad menor ni preliminar. Produce las restricciones ante las que las teorías responden, y de vez en cuando produce un número que permanece sin explicar durante cincuenta años y que durante todo ese tiempo era correcto.

En qué punto estamos

El artículo termina en preguntas que plantear a cualquier número ofrecido como medida. Qué definición está en vigor, qué validó el instrumento y qué validó eso, si la magnitud se teorizó antes, si la cifra es un valor o un residuo, si se sigue midiendo, y qué tendría que ocurrir para que cambiara. Qué establecen las secciones, y cómo, es lo que sigue.

Lo que sabemos y cómo lo sabemos

La definición permisiva toma la medida como la asignación de numerales a objetos o sucesos de acuerdo con reglas, y separa cuatro tipos de escala por la operación empírica que cada uno permite. La objeción restrictiva es que, bajo esa visión, cualquier regla define ya al menos una escala nominal, de modo que nunca se somete a prueba si el atributo es cuantitativo. Las leyes de radiación en competencia diferían solo en longitudes de onda largas; un experimento de rayos residuales con cuatro cavidades encontró que ninguna reproducía los datos a todas las temperaturas; una comunicación sobre el límite en longitudes de onda largas precedió a la ecuación que funcionó, y los dos resultados se presentaron en la misma sesión. La corrección de Le Verrier al movimiento secular del perihelio de Mercurio la recoge Laskar como unos treinta y ocho segundos de arco por siglo a partir de veintiuna observaciones de tránsitos, frente a unos cuarenta y tres de la teoría posterior. La Resolución 1 define el kilogramo fijando la constante de Planck, las siete constantes definitorias no tienen incertidumbre, y el Apéndice 2 de esa misma resolución sitúa el prototipo en 1 kg dentro de la incertidumbre típica relativa heredada, determinada experimentalmente desde entonces. El estudio de Chang se usa solo como lo describe su editorial; su cuerpo no se leyó aquí.

Lo que todavía no sabemos

Si la circularidad descrita vale para los patrones de medida en general, cosa que la sección 3 declina establecer. Cuánto tiempo se midió la constante de Planck antes de que su valor quedara fijado, y cuánto se midió la temperatura antes de que la termodinámica diera cuenta de ella, dos periodos a los que este artículo no pone cifra. Qué argumenta el libro de Chang, frente a lo que su editorial dice que examina. Y qué causó qué en octubre de 1900: el orden de los hechos queda registrado y la pregunta causal queda abierta.

Qué podría hacernos cambiar de opinión

Una lectura del argumento de Chang, y no de la descripción de su editorial, zanjaría qué establece ese estudio sobre romper el círculo. Mostrar que algún patrón de medida se validó sin apoyarse nunca en instrumentos que después juzgó estrecharía lo que la sección 3 declina generalizar. Sobre el residuo, importaría una teoría sucesora que cubriera el mismo dominio observacional y no reprodujera la cifra con la precisión pertinente; una que cubriera otro terreno no. Y una decisión metrológica que dejara de fijar una constante definitoria la devolvería a la medida.

Si solo recuerdas una cosa

Los números llegan a menudo antes que las teorías que los explican, así que pregunta ante qué respondía una medida antes de preguntar si es verdadera.

Artículos relacionados

- KNO-010 · *La clasificación como herramienta de conocimiento*

Referencias

Los números entre corchetes del cuerpo son números de afirmación y no de referencia. Cada uno remite a una fila de KNO-011_CLAIMS_AND_SOURCES_RECONSTRUCTION.md, que da una fuente académica u oficial para cada una y nombra la entrada de abajo que la sostiene. Las afirmaciones inciertas, los límites de alcance declarados y las fuentes excluidas se enumeran en KNO-011_UNCERTAINTIES_RECONSTRUCTION.md.

BIPM. 2018. *Resolution 1 of the 26th CGPM (2018): On the revision of the International System of Units (SI)*. Conférence Générale des Poids et Mesures. URL estable: <https://www.bipm.org/en/committees/cg/cgpm/26-2018/resolution-1> DOI 10.59161/CGPM2018RES1E

BIPM. 2021. *SI Brochure – 9th edition (2019) – Appendix 2*, 7 de julio de 2021. Bureau International des Poids et Mesures. URL estable: <https://www.bipm.org/documents/20126/41489673/SI-App2-kilogram.pdf>

BIPM. Página «SI defining constants». Bureau International des Poids et Mesures. **Sin fecha mostrada en la página**. URL estable: <https://www.bipm.org/en/measurement-units/si-defining-constants>

Chang, Hasok. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford Studies in Philosophy of Science. Oxford University Press. ISBN impreso 9780195171273. URL estable: <https://global.oup.com/academic/product/inventing-temperature-9780195171273> **Manifestación en línea, Oxford Academic**: ISBN en línea 9780199835003, DOI 10.1093/0195171276.001.0001, <https://academic.oup.com/book/5530>

González de Arrieta, Iñigo. 2022. «Beyond the infrared: a centenary of Heinrich Rubens's death». *The European Physical Journal H* 47: artículo 11. DOI 10.1140/epjh/s13129-022-00044-x

Laskar, Jacques. 2017. «Des premiers travaux de Le Verrier à la découverte de Neptune». *Comptes Rendus Physique* 18: 504–519. DOI 10.1016/j.crhy.2017.10.011

Michell, Joel. 1997. «Quantitative science and the definition of measurement in psychology». *British Journal of Psychology* 88: 355–383. DOI 10.1111/j.2044-8295.1997.tb02641.x URL estable: <https://errorstatistics.com/wp-content/uploads/2015/02/michell-1997-1.pdf>

Stevens, S. S. 1946. «On the Theory of Scales of Measurement». *Science* 103 (2684): 677–680. **Sin DOI localizado sobre página consultada**. URL estable: <https://dl.icdst.org/pdfs/files3/38396a3ee53de9478a57af4605f0e1c1.pdf>

KNO-012 · Experimentos y control de variables

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

En los diseños comparativos con asignación aquí tratados, un experimento es una comparación que alguien dispuso, y de esa disposición vienen su ventaja sobre mirar y las dificultades que siguen. Muchos experimentos no asignan nada a nadie, y nada de lo que sigue define el experimento en general. Este artículo sigue la disposición y no el resultado. Separa las dos tareas que se le piden a la palabra aleatorización, generar una secuencia impredecible y ocultar esa secuencia a quien recluta participantes, y registra el precio medido de fallar en la segunda. Distingue las protecciones que terminan en la asignación de las que deben continuar después, razón por la que el enmascaramiento se evalúa aparte. Pasa a los casos en que la intervención es imposible y la comparación hay que encontrarla. Cierra donde el resultado mismo se discute y juzgar la competencia del trabajo exige saber cuál debía ser la respuesta.

En 60 segundos

En un diseño comparativo con asignación, un experimento compra una sola cosa: el investigador decide quién recibe qué, de modo que la razón por la que un grupo recibió el tratamiento tiene respuesta conocida en lugar de desconocida. Los experimentos sin asignación quedan fuera de lo que sigue. Todo lo demás del método protege esa propiedad. La decisión misma es por tanto el punto vulnerable, y la palabra aleatorización cubre ahí dos tareas distintas: generar una secuencia que nadie pueda predecir y ocultar esa secuencia a quien recluta participantes. Las normas de publicación las numeran por separado, y los ensayos juzgados con ocultación inadecuada declaran efectos mayores que los juzgados con ocultación adecuada. La ocultación se detiene en el momento de la asignación. El enmascaramiento es otro instrumento, no siempre está disponible, e importa más quién quedó enmascarado que si aparece la palabra. Donde nadie puede asignar nada, la comparación hay que encontrarla, y afirmar que su variación es como si fuera aleatoria es un supuesto sobre el mundo y no un hecho sobre el diseño. Lo más difícil es un resultado en disputa, donde los juicios sobre competencia y sobre la respuesta se mueven juntos.

Por qué importa esta pregunta

Las protecciones que aquí se tratan pueden ser la diferencia entre una comparación cuya garantía es defendible y otra cuyo resultado aparente desborda su diseño. Y son además invisibles en el resultado: un ensayo declara un tamaño de efecto igual tanto si su secuencia estuvo oculta como si no, y la cifra que sale no lleva marca alguna de la disposición que la produjo. Leer bien un experimento consiste por eso en leer lo que se hizo antes de que los datos existieran, que es la parte que un resultado publicado menos suele mostrar y de la que más suele depender.

1. Lo que un experimento añade

La observación registra lo que ocurrió. Un experimento dispone antes la situación y registra después lo que ocurrió, y la disposición es toda la diferencia.

Esa diferencia compra una sola cosa. **En un experimento es el investigador quien decide quién recibe qué**, de modo que la pregunta de por qué un grupo recibió el tratamiento y otro no tiene una respuesta conocida en lugar de una desconocida. Todo lo demás del método experimental existe para proteger esa única propiedad, y cada caso de las secciones siguientes es un caso de protegerla, de perderla o de trabajar sin ella.

Qué cubre este artículo y qué no. Muchos experimentos no asignan nada a nadie: varían un voltaje, una temperatura, un reactivo o un retardo, sin participantes, sin grupos y sin aleatorización. Lo que sigue trata de **diseños comparativos en los que se asignan unidades a condiciones**, porque esa es la familia en la que las protecciones aquí discutidas hacen su trabajo. Nada de ello debe leerse como una definición de experimento en general.

La palabra control hace aquí dos trabajos y conviene separarlos desde el principio. **Mantener fija una variable** es disponer que no varíe a lo largo de la comparación. **Un grupo de control** es el grupo de comparación frente al que se juzga la intervención: puede recibir un placebo, la atención habitual, otra intervención o nada en absoluto, y lo que lo hace control es el papel que desempeña en la comparación y no la ausencia de tratamiento en él. Lo primero es una propiedad del montaje; lo segundo es una parte del diseño. Un estudio puede tener grupo de control y no mantener nada fijo, y puede mantener mucho fijo y no tener nada con qué comparar.

Ningún montaje mantiene todo fijo. La igualdad completa de condiciones a lo largo de una comparación no es alcanzable, y por eso el argumento a favor de una comparación versa siempre sobre qué diferencias se controlaron y cuáles se dejaron al azar, y nunca sobre su ausencia. **Este es un enunciado de encuadre, declarado como tal en la sección 4 de KNO-012_UNCERTAINTIES_RECONSTRUCTION.md.**

En qué punto estamos

La observación registra lo que ocurrió; un experimento dispone antes la situación, y la disposición es toda la diferencia. En los diseños comparativos con asignación tratados aquí compra una propiedad: el investigador decide quién recibe qué, y el resto del método la protege. Se separan dos sentidos de control: mantener fija una variable pertenece al montaje; un grupo de control, al diseño. Un estudio puede tener uno sin el otro.

2. Por qué la asignación es el paso que soporta la carga

Si el investigador decide quién recibe qué, la decisión misma se convierte en el punto vulnerable.

Fisher dio la respuesta clásica, y le puso un título que enuncia el argumento entero: la sección 9 de *The Design of Experiments* se titula «Randomisation; the Physical Basis of the Validity of the Test» [1]. Su afirmación allí es que no hace falta más aparato, porque la simple precaución de la aleatorización basta para garantizar la validez de la prueba de significación [2].

Dos rasgos de esa frase merecen marcarse, y un tercero pertenece a su contexto. Es una afirmación sobre la **validez de una prueba**, no sobre el equilibrio de ningún experimento concreto, y este artículo no la extiende más allá de lo que dice. Es además condicional y no incondicional: la garantía se ofrece

sobre un fondo en el que los demás errores evitables se han evitado, no como un aval que sobreviva a ellos. **La afirmación se trata aquí como condicionada por el contexto experimental circundante; no se interpreta como si la aleatorización eliminara todas las demás fuentes evitables de error.** Y describe la aleatorización como una **precaución**, que es la categoría correcta: es un procedimiento adoptado de antemano, cuyo valor no depende del aspecto que los grupos resultantes tengan después.

En qué punto estamos

El punto vulnerable es la decisión misma, en cuanto alguien la toma. La respuesta de Fisher atañe a la validez de una prueba y no al equilibrio de un experimento cualquiera. El artículo trata la afirmación como condicionada por un contexto en el que se evitaron otros errores evitables, y no como una garantía de que la aleatorización los elimine. Llama precaución a la aleatorización.

3. Dos tareas, una palabra

En la práctica se le piden a la aleatorización dos cosas distintas, y solo una de ellas es aritmética.

La primera es **generar** una secuencia de asignación que nadie pueda predecir. La segunda es **ocultar** esa secuencia a quien recluta a los participantes, de modo que saber a qué rama caerá la siguiente persona no pueda influir en si esa persona llega a ser reclutada. Las normas de publicación tratan estas obligaciones por separado: la lista de comprobación CONSORT 2010 recoge la generación de la secuencia, el mecanismo de ocultación de la asignación y el enmascaramiento como tres puntos numerados por separado, 8a, 9 y 11 [3].

La separación no es pulcritud burocrática, y lleva un número asociado. Schulz y colaboradores compararon ensayos según lo adecuado de su ocultación y hallaron que las razones de momios quedaron exageradas en un 41 % en los ensayos con ocultación inadecuada y en un 30 % en los de ocultación poco clara, ajustando por otros aspectos de la calidad [4].

Ese resultado hay que leerlo con cuidado, porque es fácil exagerarlo. Se refiere al conjunto de ensayos examinados en aquel estudio, es una asociación entre un rasgo metodológico declarado y el tamaño de los efectos declarados, y la exageración se mide frente a ensayos juzgados con ocultación adecuada, no frente a una verdad conocida. Lo que establece es estrecho y grave: **una aleatorización bien generada y mal ocultada no queda protegida por haber sido bien generada.** Lo que no establece es nada sobre la otra mitad del par. La comparación ordena los ensayos según lo adecuado de su ocultación, de modo que no muestra que las secuencias estuvieran correctamente generadas en aquellos cuya ocultación se juzgó inadecuada. Solo una de las dos propiedades es aquella por la que este estudio ordena.

En qué punto estamos

Una palabra cubre dos tareas. Generar una secuencia que nadie pueda predecir es aritmética; ocultarla a quien recluta participantes no lo es, y las normas de publicación las numeran aparte. El precio medido de fallar en la segunda es una asociación dentro de un conjunto de ensayos, medida frente a los juzgados con ocultación adecuada y no frente a una verdad conocida. Nada dice de la otra mitad del par.

4. Antes de la asignación y después

La asignación protege un momento. Casi todo el experimento ocurre después de ese momento, y para eso hacen falta instrumentos distintos.

El manual Cochrane traza la línea de forma explícita. La ocultación de la secuencia de asignación busca evitar el sesgo en la asignación de la intervención impidiendo que el personal del ensayo y los participantes conozcan la secuencia antes de la asignación y hasta ella [5], y, de forma decisiva, puede implementarse siempre con éxito, con independencia del diseño del estudio o del área clínica [6]. El enmascaramiento es un instrumento de la clase contraria: busca evitar el sesgo después de la asignación y no siempre puede implementarse [7].

Esas dos frases juntas son lo más útil de este artículo. Una protección está disponible en los diseños y áreas clínicas que el manual describe, que son los ensayos aleatorizados, y tiene un límite temporal. Leída fuera de ese marco, la palabra siempre afirmarí más de lo que dice la fuente. La otra no siempre está disponible, porque hay intervenciones que no pueden disimularse, y donde lo está no tiene por qué extenderse a toda la duración del ensayo: ha de sostenerse en las personas, y en las etapas, en que conocer la asignación pudiera introducir sesgo. Un ensayo puede por tanto ser impecable en el punto de la asignación y quedar expuesto durante el resto de su duración, y los dos estados deben evaluarse por separado en lugar de resumirse en una única puntuación de calidad.

El mismo manual va más lejos y separa al evaluador del participante: el enmascaramiento de quienes evalúan el resultado, para evitar el sesgo en la *medición* del resultado, se considera aparte, en el dominio de sesgo en la medición de los resultados [8]. **Importa más quién quedó enmascarado que si aparece la palabra enmascarado.** Un ensayo en el que los participantes no pudieron enmascararse pero sí los evaluadores, y otro en el que ocurre lo contrario, están expuestos a fallos distintos y no deberían describirse con la misma expresión.

En qué punto estamos

La asignación protege un momento y casi todo el ensayo ocurre después. La ocultación evita el sesgo antes de la asignación y hasta ella y, dentro de los ensayos aleatorizados que el manual describe, puede implementarse siempre. El enmascaramiento actúa después y no siempre puede implementarse, y donde puede no tiene por qué cubrir toda la duración. Importa por eso más quién quedó enmascarado que si aparece la palabra.

5. Cuando la comparación hay que encontrarla y no fabricarla

Para muchas preguntas la disposición de la sección 1 sencillamente no está disponible. El documento de fundamentación del comité del premio enuncia la restricción sin rodeos: por limitaciones económicas, éticas o prácticas, muchas preguntas importantes, en particular en las ciencias sociales, no pueden estudiarse mediante un experimento aleatorizado controlado [9].

La respuesta consiste en buscar situaciones que hicieran la disposición por uno. El mismo documento define el objeto: un experimento natural es un suceso o una situación, no sometida al control de los sujetos estudiados, que genera variación en la variable de interés como si hubiera sido asignada al azar [10].

La carga de esa definición la llevan dos expresiones y ambas son condiciones, no descripciones. El suceso debe quedar fuera del control de los sujetos, lo que cierra una vía por la que las personas estudiadas podrían seleccionarse a sí mismas hacia el grupo que les conviene. No cierra

por sí solo las demás. La selección ejercida por alguien distinto de los sujetos, la anticipación del suceso y las diferencias sistemáticas que el suceso acompaña en lugar de crear sobreviven todas a esa condición y hay que argumentarlas caso por caso. Que la variación resultante sea comparable a una asignación al azar es por tanto una afirmación que exige justificación, y el instrumental para evaluar esas justificaciones es la inferencia causal, que esta serie aborda en KNO-014 y no aquí. Y la variación debe ser *como si* estuviera asignada al azar, lo cual es un supuesto sobre el mundo y no un hecho sobre el diseño, porque nadie aleatorizó nada.

Esta es la diferencia con la sección 2 y no debe difuminarse. En un experimento aleatorizado la razón de la asignación se conoce porque alguien la produjo. En un experimento natural la razón se argumenta. Ese argumento puede ser fuerte, y el diseño puede ser la única vía disponible hacia una pregunta que importa, pero es una garantía de otra clase y este artículo no trata las dos como intercambiables.

En qué punto estamos

Donde la disposición no está disponible, la comparación hay que encontrarla. La definición lleva dos condiciones y no dos descripciones: el suceso queda fuera del control de los sujetos, lo que cierra una vía de autoselección y no las demás, y la variación es *como si* estuviera asignada al azar, que es un supuesto sobre el mundo. En un diseño aleatorizado la razón se conoce; aquí se argumenta.

6. Cuando lo que está en disputa es el resultado mismo

El caso más difícil no es un experimento mal hecho. Es un experimento bien hecho cuyo resultado se discute, porque entonces la cuestión de si estuvo bien hecho no puede zanjarse con independencia del resultado.

Collins formuló la circularidad de la manera que le dio nombre, en un pasaje que Franklin cita de él: no tendremos idea de si podemos hacerlo hasta que intentemos ver si obtenemos el resultado correcto; pero ¿cuál es el resultado correcto?; cuál sea el resultado correcto depende de si hay o no ondas gravitatorias alcanzando la Tierra en flujos detectables; para averiguarlo hemos de construir un buen detector de ondas gravitatorias y mirar; pero no sabremos si hemos construido un buen detector hasta que lo hayamos probado y hayamos obtenido el resultado correcto [11].

Ese es el diagnóstico de Collins, y aquí queda registrado como suyo y no adoptado. Según él, la competencia se juzga por el resultado mientras que el resultado se acepta según la competencia. En un campo asentado esto es invisible, porque todo el mundo sabe ya qué produce un instrumento que funciona. Se vuelve visible precisamente cuando un campo no está asentado, que es cuando importa.

De ahí no se sigue que el círculo no pueda romperse, y este artículo no afirma que no pueda.

Franklin, que recoge el pasaje, argumenta en contra de esa conclusión y sostiene que el regreso puede romperse mediante argumentación razonada [12]. Las dos posiciones quedan registradas y no se adopta ninguna, y conviene notar que no son simétricas: una dice que existe una dificultad, la otra que es superable, y ambas pueden ser ciertas.

El límite de alcance fijado antes en esta serie se conserva aquí. El regreso es un problema real en algunos casos disputados, no una ley general según la cual toda calibración sea circular. Un termómetro contrastado con un patrón nacional no está en un regreso, y describirlo como si lo estuviera vaciaría el concepto del contenido que lo hace útil.

En qué punto estamos

Un experimento bien hecho cuyo resultado se discute es el caso difícil, porque entonces la competencia no puede juzgarse con independencia del resultado. Collins formula ese círculo y Franklin, que lo cita, sostiene que puede romperse mediante argumentación razonada. Ambas posiciones quedan registradas sin adoptar ninguna, y no son simétricas. El regreso pertenece a los casos disputados y no a la calibración en general.

7. Leer bien un experimento

¿Se asignó algo, o solo se observó? Si nadie asignó, la razón por la que cada persona está en el grupo en que está hay que argumentarla en lugar de conocerla.

¿Quién sabía qué, y cuándo? Generación de la secuencia y ocultación de la asignación son fallos distintos. También lo son ocultación y enmascaramiento, y también el enmascaramiento de los participantes y el de quienes evalúan el resultado.

¿Qué protecciones terminaron en la asignación? La ocultación está disponible en los ensayos aleatorizados que el manual describe, y se detiene en la asignación. El enmascaramiento no siempre está disponible, y donde lo está, ha de sostenerse en las personas y etapas en que conocer la asignación pudiera sesgar lo que ocurre o lo que se registra.

¿Cuál es el grupo de comparación, y qué lo hace comparable? En un diseño aleatorizado la respuesta es procedimental. En un experimento natural es un argumento sobre el mundo, y el argumento debería estar a la vista.

¿Está en disputa el resultado? Si lo está, cabe esperar que los juicios sobre la competencia y los juicios sobre la respuesta se muevan juntos, y conviene buscar las razones ofrecidas en lugar de un experimento decisivo.

¿Qué se mantuvo fijo, y era eso lo mismo que tener un grupo de control? Son cosas separadas, y un estudio puede tener una sin la otra.

Un experimento no es un dispositivo que produce verdad. Es una disposición que sustituye una incógnita, por qué estas personas y no aquellas, por un procedimiento, y que dedica el resto de su duración a defender esa sustitución.

En qué punto estamos

El artículo termina en preguntas para llevarse a un experimento: si se asignó algo, quién sabía qué y cuándo, qué protecciones se detuvieron en la asignación, qué hace comparable al grupo de comparación, si el resultado está en disputa y si lo que se mantuvo fijo era lo mismo que un grupo de control. Una disposición, no un dispositivo que produce verdad. Qué sabemos, y cómo, es lo que sigue.

Lo que sabemos y cómo lo sabemos

Fisher tituló la sección 9 de *The Design of Experiments* «Randomisation; the Physical Basis of the Validity of the Test», y sostuvo allí que la simple precaución de la aleatorización basta para garantizar la validez de la prueba de significación. Es una afirmación sobre la validez de una prueba y no sobre el

equilibrio de ningún experimento concreto, y el contexto deja claro además que la validez no equivale a afirmar que la aleatorización elimine todas las demás fuentes evitables de error. La lista CONSORT 2010 recoge la generación de la secuencia, el mecanismo de ocultación y el enmascaramiento como tres puntos numerados por separado. Schulz y colaboradores, comparando ensayos según lo adecuado de su ocultación, informan de razones de momios exageradas en un 41 % con ocultación inadecuada y en un 30 % con ocultación poco clara. El manual Cochrane dice que la ocultación evita el sesgo antes de la asignación y hasta ella y, dentro de los ensayos aleatorizados que el manual describe, puede implementarse siempre, que el enmascaramiento lo evita después y no siempre puede, y que el de quienes evalúan el resultado se considera aparte. El documento del comité del premio enuncia que muchas preguntas importantes no pueden estudiarse con experimento aleatorizado controlado, y define el experimento natural. Collins formula el regreso en un pasaje que Franklin cita, y Franklin sostiene que puede romperse mediante argumentación razonada.

Lo que todavía no sabemos

Si las secuencias estaban correctamente generadas en los ensayos cuya ocultación se juzgó inadecuada, ya que la comparación de Schulz ordena por ocultación y no por generación. Hasta dónde viajan las cifras de exageración más allá del conjunto de ensayos examinados, siendo una asociación entre un rasgo declarado y los tamaños de efecto declarados y no una distancia respecto de una verdad conocida. Si la variación de un experimento natural concreto es de veras como si estuviera asignada al azar, que es un supuesto sobre el mundo y no un hecho sobre el diseño. Y si el regreso puede romperse en un caso disputado dado, donde Collins y Franklin quedan registrados sin que se adopte ninguno.

Qué podría hacernos cambiar de opinión

Una comparación que ordenara los ensayos según lo adecuado de la generación de la secuencia y no de la ocultación diría algo sobre la mitad del par que esta evidencia deja intacta. Estudios de ocultación fuera de los ensayos aleatorizados que el manual Cochrane describe pondrían a prueba hasta dónde llega ese siempre. Para un experimento natural, evidencia que cerrara las vías que la definición deja abiertas, la selección ejercida por alguien distinto de los sujetos, la anticipación del suceso y las diferencias sistemáticas que el suceso acompaña, acercaría el argumento a la garantía procedimental que sustituye.

Si solo recuerdas una cosa

En los diseños comparativos con asignación aquí tratados, la asignación sustituye la razón desconocida de por qué estas unidades y no aquellas por un procedimiento, y el resto la defiende.

Artículos relacionados

- KNO-014 · *Causalidad y correlación*

Referencias

Los números entre corchetes del cuerpo son números de afirmación y no de referencia. Cada uno remite a una fila del registro de afirmaciones, que nombra la fuente de abajo que la sostiene.

Comité del Premio en Ciencias Económicas. 2021. *Answering Causal Questions Using Observational Data. Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2021*. The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel, The Royal Swedish Academy of Sciences, 11 de octubre de 2021. URL estable: <https://www.nobelprize.org/uploads/2021/10/advanced-economicsciencesprize2021.pdf>

Fisher, R. A. 1935. *The Design of Experiments*. Edimburgo y Londres: Oliver and Boyd. **Sin ISBN: la obra es anterior al sistema ISBN**. URL estable del ejemplar digitalizado consultado: <https://archive.org/details/in.ernet.dli.2015.502684>

Franklin, Allan. 1994. «How to Avoid the Experimenters' Regress». *Studies in History and Philosophy of Science* 25 (3): 463–491. **Sin DOI mostrado sobre página consultada; identificador impreso 0039-3681/94**. URL estable: https://physics.yale.edu/sites/default/files/files/3_1%20franklin.pdf

Higgins, Julian PT, Jelena Savović, Matthew J Page, Roy G Elbers, y Jonathan AC Sterne. 2024. «Chapter 8: Assessing risk of bias in a randomized trial». En *Cochrane Handbook for Systematic Reviews of Interventions*, versión 6.5. Cochrane. URL estable: <https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-08>

Schulz, Kenneth F., Douglas G. Altman, y David Moher, for the CONSORT Group. 2010. «CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials». *PLoS Medicine* 7 (3): e1000251. DOI 10.1371/journal.pmed.1000251

Schulz, Kenneth F., Iain Chalmers, Richard J. Hayes, y Douglas G. Altman. 1995. «Empirical Evidence of Bias: Dimensions of Methodological Quality Associated With Estimates of Treatment Effects in Controlled Trials». *JAMA* 273 (5): 408–412. DOI 10.1001/jama.1995.03520290060030 URL estable: https://www.jameslindlibrary.org/wp-data/uploads/2012/02/Schulz_Chalmers_Hayes_Altman_1995.pdf

KNO-013 · Modelos, mapas y realidad

KNO · Formas de conocer

Juan Segurola

v0.1 · ES

Todos los derechos reservados

Resumen del artículo

Un modelo científico no es una réplica. Está en lugar de una parte o un aspecto seleccionados del mundo, y seleccionar es el trabajo. Este artículo trata la omisión y la distorsión deliberada como piezas de funcionamiento y no como defectos, y separa el error, que quien modeliza corregiría, de la idealización, que defendería. Registra una posición publicada según la cual las leyes que mejor unifican son las que menos describen, sin adoptarla. Enuncia la restricción que un biólogo de poblaciones formuló en 1966 y la objeción que niega que esa restricción sea necesaria. Muestra que la frase más citada sobre modelos no está en el artículo que suele citarse. Cierra en la cuestión del realismo que dos artículos anteriores entregaron, enuncia los dos criterios y no adopta ninguno.

En 60 segundos

Un modelo representa una parte elegida del mundo y no el mundo, así que preguntar si es bueno es preguntar para qué se construyó. Dejar algo fuera no es lo mismo que equivocarse: una abstracción omite, una idealización distorsiona a propósito, y solo la segunda afirma una falsedad deliberadamente. El criterio es la intención, porque lo que quien modeliza corregiría al advertirlo es un error y lo que defendería tiene que ganarse su sitio. Una posición desarrollada sostiene que las leyes más unificadoras son las menos descriptivas, y aquí se registra sin adoptarse. Una restricción de 1966 dice que generalidad, detalle conservado y precisión no pueden maximizarse a la vez en un modelo manejable; una objeción niega que ese intercambio sea necesario, no que ocurra. La frase famosa sobre que los modelos son erróneos está en un documento distinto del que suele citarse. Si las partes de un modelo son reales queda abierto, pero el ajuste nunca se acredita solo.

Por qué importa esta pregunta

Un modelo en uso no es una copia pequeña de nada, y la pregunta útil sobre él no es si es verdadero sino para qué se construyó y a qué renunció al hacerlo. Sin eso, un modelo se descarta por las partes que nunca reclamó o se le confía un alcance que nunca tuvo. Las distinciones que aquí se hacen son las que permiten a quien lee preguntar qué falsedad se eligió, qué compró, y si una formulación memorable está soportando una carga que su fuente no sostiene.

1. Un modelo no es una copia pequeña

La palabra modelo sugiere una réplica a escala, y la sugerencia induce a error de una manera concreta. Una réplica pretende conservarlo todo a tamaño reducido. Un modelo científico pretende algo más estrecho: sustituye a una parte elegida del mundo, y la elección es el trabajo.

La formulación de referencia pone el objetivo por delante. Muchos modelos científicos son modelos representacionales, y representan una parte o un aspecto seleccionados del mundo, que es el sistema objetivo del modelo [1]. **La expresión que hay que retener es parte o aspecto seleccionados.** Un modelo no representa el mundo; representa un trozo de él, elegido de antemano, y la pregunta de si un modelo es bueno es siempre la pregunta de si es bueno para ese trozo y para ese propósito.

Una advertencia sobre la analogía del mapa, que este artículo usa con parquedad y a propósito. Los mapas son la ilustración permanente de la representación selectiva, y un artículo anterior de esta serie los trata de forma directa. Pero la restricción que ata a los mapas con más fuerza es un teorema geométrico sobre proyectar una esfera en un plano, y **aquí no se establece ningún teorema correspondiente sobre los modelos en general, ni se transporta el de KNO-006.** Tomar prestada la fuerza del mapa sin tomar prestada su matemática produciría una frase impresionante y sin apoyo. Este artículo usa por tanto el mapa como ilustración de la selección y nunca como prueba sobre lo que los modelos pueden o no pueden hacer.

En qué punto estamos

La palabra modelo sugiere una réplica, y ahí está lo engañoso. Un modelo científico está en lugar de una porción elegida del mundo, y elegirla es el trabajo, de modo que si es bueno es si lo es para esa porción y ese propósito. Los mapas ilustran aquí esa elección y nunca son prueba, porque su restricción propia es un teorema que este artículo no traslada.

2. Dejar cosas fuera no es lo mismo que equivocarse

Los modelos no son una sola clase de cosa, y las distinciones importan para lo que sigue. Un modelo **representacional** sustituye a un sistema objetivo. Una **abstracción** deja fuera rasgos sin afirmar que estén ausentes. Una **idealización** simplifica o distorsiona a propósito lo que conserva. Solo la tercera afirma algo falso deliberadamente, y si un modelo solo omitiera sería un compendio y no una distorsión.

La distinción que importa es entre un error y una idealización. **Un error es algo que quien modeliza retiraría al descubrirlo. Una idealización es algo que quien modeliza puso ahí y defendería.** La formulación de referencia es explícita sobre la intención: los modelos idealizados implican una simplificación o distorsión deliberada de algo complicado, con el objetivo de hacerlo más tratable o más comprensible [2].

Esa definición contiene un criterio aplicable. **Una idealización responde ante un propósito.** Un plano sin rozamiento, una población bien mezclada, un sólido perfectamente rígido: cada uno es falso, cada uno está elegido, y a cada uno puede preguntársele si se gana su falsedad en el uso al que se destina el modelo. Una distorsión que ha dejado de ganársela se ha convertido en un error sin que nadie haya cambiado las ecuaciones.

En qué punto estamos

Aquí se separan tres cosas. Un modelo representacional está en lugar de un objetivo; una abstracción omite sin afirmar ausencia; una idealización distorsiona a sabiendas, y solo ella afirma una falsedad a propósito. El criterio es de intención y no de exactitud: lo que quien modeliza tacharía al advertirlo es un error, y lo que defendería tiene que ganarse su falsedad frente al uso.

3. Un argumento publicado que va más lejos

Hay en la literatura una posición más fuerte, y conviene enunciarla como una posición y no como el consenso.

Cartwright sostiene que las leyes explicativas fundamentales, los éxitos más profundos y admirados de la física moderna, no describen de hecho regularidades que existan en la naturaleza [3]. El argumento no es que la física fracase. Es que el poder explicativo y la exactitud descriptiva se obtienen por medios distintos y se intercambian entre sí, de modo que las leyes que más unifican son las que menos describen.

El mismo argumento lleva una distinción que le impide colapsar en escepticismo general: según esa exposición, las entidades teóricas y las leyes complejas y localizadas que las describen pueden interpretarse de forma realista, pero las leyes unificadoras simples de la teoría básica no [4]. **El escepticismo apunta a la cúspide de la jerarquía y no a su base.**

Este artículo no adopta esa posición. La registra porque es un argumento desarrollado en la literatura y porque afila la pregunta que la sección 6 tiene que responder. Conviene que el lector sepa que está en discusión, y que las frases citadas aquí proceden de la descripción editorial del libro y no de una lectura de su argumento, limitación que consta en el registro de incertidumbres.

En qué punto estamos

Existe en la literatura una posición más fuerte y este artículo la registra sin adoptarla. Según ella las leyes explicativas más profundas no describen regularidades, porque alcance y exactitud se compran por vías distintas, y su propia distinción la libra del escepticismo general: las leyes localizadas y las entidades que describen pueden leerse en clave realista, las unificadoras no. Las frases citadas vienen de una descripción editorial.

4. El intercambio que no puede sortearse con diseño

Uno de los resultados más útiles sobre modelización es una restricción, y quedó enunciado con claridad en 1966. **Es una posición en un debate vivo y no un teorema zanjado**, y la objeción queda recogida más abajo y no remitida al registro de incertidumbres.

Levins escribió que por supuesto es deseable trabajar con modelos manejables que maximicen generalidad, realismo y precisión hacia los objetivos solapados pero no idénticos de comprender, predecir y modificar la naturaleza, pero que esto no puede hacerse [5].

Conviene leer la frase con cuidado, porque su estructura suele perderse. **Nombra tres propiedades, las asocia a tres objetivos distintos y dice que las tres no pueden maximizarse a la vez en un modelo manejable.** No dice que ninguna de ellas sea inalcanzable, y no las jerarquiza. **La palabra realismo no significa aquí lo que en la sección 6:** en la tríada de Levins designa cuánto detalle de un sistema conserva un modelo, mientras que en la sección 6 es una posición sobre si debe creerse el contenido inobservable de una teoría. Son cosas distintas y este artículo no deja que una arrastre a la otra. Lo que dice es que en un modelo manejable las tres no pueden maximizarse a la vez, de modo que donde se fuerza una las otras quedan bajo presión, y que ese intercambio es una decisión se advierta o no al tomarla. **Es la afirmación de Levins sobre la construcción de modelos tal como él la formuló, y este artículo no la extiende a una ley que todo modelo deba obedecer.**

La objeción es directa y no es una discusión sobre palabras. Orzack y Sober rechazan que generalidad, realismo y precisión tengan que intercambiarse necesariamente entre sí, de modo que el conflicto que describe Levins no es una restricción universal sobre la construcción de modelos [12]. Conviene leerla tan estrechamente como está escrita: niega que el intercambio sea necesario, no que alguna vez ocurra. Quien encuentre una de las tres bajo presión no queda refutado, y quien las encuentre mejorando a la vez no ha roto ninguna ley. **Lo que la objeción retira es el derecho a tratar la restricción como zanjada**, que es justo lo que esta sección ya se negaba a reclamar.

Este es el contenido práctico de todo lo anterior. La selección, la idealización y la distancia entre explicar y describir aparecen aquí como un único presupuesto. Un modelo muy general difícilmente será muy preciso sobre un caso particular. Un modelo afinado en detalle para un caso difícilmente viajará. **Preguntar a cuál de las tres renunció un modelo suele ser más informativo que preguntar si el modelo es correcto.**

En qué punto estamos

La restricción dice que generalidad, detalle conservado y precisión no pueden maximizarse a la vez en un modelo manejable, y ese detalle no es la doctrina de la sección 6. La objeción recogida al lado niega que el intercambio sea necesario, no que ocurra, así que ni la presión sobre una propiedad ni la mejora de las tres zanján nada. Queda cuál se gastó, y en qué.

5. Equivocado, y útil, en dos documentos distintos

Una frase sobre modelos se cita mucho más que el resto de la literatura sobre ellos, y a menudo se atribuye al documento equivocado. **Con qué frecuencia no se establece aquí:** no se ha consultado ningún estudio bibliométrico, y lo que la sección sostiene no lo necesita.

El artículo de 1976 es el que suele citarse. Lo que contiene en realidad es esto: puesto que todos los modelos son erróneos, el científico no puede obtener uno correcto por elaboración excesiva; al contrario, siguiendo a Guillermo de Occam, debe buscar una descripción económica de los fenómenos naturales [6]. Eso es un argumento completo por sí solo, y no es la frase célebre. **Una comprobación directa de ese artículo confirma que la expresión pero algunos son útiles no aparece en ninguna parte de él [7].**

La redacción célebre está en un informe técnico de 1979, donde aparece la cadena literal «ALL MODELS ARE WRONG BUT SOME ARE USEFUL» [8], acompañada de la observación de que sería muy notable que algún sistema existente en el mundo real pudiera representarse exactamente mediante un modelo simple [9].

Los dos textos dicen cosas compatibles y no son el mismo texto. El pasaje de 1976 es un argumento contra la elaboración excesiva: puesto que ningún modelo es correcto, añadir parámetros no puede volverlo correcto, de modo que la economía es la respuesta adecuada. La redacción de 1979 es un argumento sobre qué preguntar: no si el modelo es verdadero, sino si es útil. Atribuir la segunda frase al primer artículo es un error pequeño con una moraleja general, y este artículo la enuncia porque la moraleja es la suya propia: **una formulación que circula por su cuenta es la que menos probablemente se comprueba contra su fuente.**

En qué punto estamos

Una frase circula más que la literatura que la rodea y suele atribuirse al documento equivocado. El artículo de 1976 sostiene que elaborar más no rescata un modelo y la economía es la respuesta, y una comprobación directa no la encuentra en él; la redacción pertenece a un informe de 1979. La moraleja es la del propio artículo: lo que viaja solo es lo que nadie comprueba.

6. La pregunta que dos artículos anteriores entregaron hacia adelante

Dos artículos anteriores de esta serie se detuvieron ante la misma puerta. Uno preguntó si un esquema de clasificación se corresponde con una división real del mundo y declinó responder. El otro preguntó si un número medido se corresponde con algo real y declinó igualmente. Ambos entregaron aquí la pregunta, y es la misma pregunta con ropa distinta.

Las dos posiciones deben atribuirse antes de formular cualquier declaración de neutralidad.

El realismo científico, tal como lo enuncia la formulación de referencia, es una actitud epistémica positiva hacia el contenido de nuestras mejores teorías y modelos, que recomienda creer tanto en los aspectos observables como en los inobservables del mundo descrito por las ciencias [10]. Según esta posición, el contenido inobservable de una teoría exitosa es candidato a la creencia, y un modelo que funciona es indicio de que sus postulados son aproximadamente correctos.

El criterio rival es la adecuación empírica, asociada al empirismo constructivo y recogida en la misma obra de referencia a través de la formulación de van Fraassen: una teoría es empíricamente adecuada exactamente si lo que dice sobre las cosas y los sucesos observables del mundo es verdadero [11]. Según esta posición, una teoría que salva los fenómenos observables ha hecho todo lo que se le exige, y creer en su maquinaria inobservable es opcional. **El criterio no equivale a ajustar un conjunto finito de datos.** Tal como se formula, alcanza a lo que la teoría dice sobre lo observable, incluido lo observable con lo que no se la ha contrastado, de modo que un buen ajuste a los datos disponibles es evidencia sobre la adecuación empírica y no la adecuación misma.

Enunciadas ambas, este artículo no adopta ninguna, y la razón no es diplomática. El desacuerdo versa sobre qué autoriza el éxito, y nada en la práctica de modelizar lo zanja: el mismo modelo, con idéntico rendimiento, es razón para creer según una posición e instrumento que funciona bien según la otra. **Lo que el desacuerdo sí zanja es un punto procedimental que ambas partes aceptan.**

Cualquiera que sea el criterio vigente, que un modelo ajuste al registro observable es ya un resultado que hay que establecer, y depende de los datos usados, de cómo se trataron, de las reglas de comparación, de los supuestos incorporados al ajuste y de la incertidumbre asociada al resultado. **El ajuste no se acredita a sí mismo.** Cualquier afirmación ulterior sobre a qué corresponden las partes del modelo es después otro paso aparte, y hay que argumentarlo aparte. **Los artículos que entregaron esta pregunta pueden por tanto seguir adelante, y lo que obtienen es menos que una respuesta.** La cuestión del realismo no queda resuelta aquí y este artículo no la resuelve. Lo que se establece es que una clasificación y una medición pueden evaluarse **operacionalmente**, por lo que hacen y para qué se construyeron, sin que esa evaluación quede a la espera de la cuestión del realismo. La cuestión sigue abierta y sigue entregada hacia adelante.

En qué punto estamos

Se exponen dos criterios y no se adopta ninguno, porque el desacuerdo versa sobre qué autoriza el éxito y modelizar no lo decide. Lo que ambas partes conceden es procedimental. Que un modelo case con el registro observable es ya un resultado por establecer, y descansa en los datos, su tratamiento, las reglas de comparación, los supuestos y la incertidumbre. La correspondencia es otro paso, y se argumenta aparte.

7. Leer bien un modelo

¿Cuál es el objetivo? Un modelo representa una parte o un aspecto seleccionados, y criticarlo por las partes que nunca reclamó es un error de categoría.

¿Qué distorsiones son deliberadas? Conviene preguntar qué falsedades defendería quien modeliza y cuáles corregiría. Las segundas son errores; las primeras tienen que ganarse su sitio.

¿A qué se renunció? Según Levins, las tres no pueden maximizarse a la vez en un modelo manejable. Identificar cuál se gastó y en qué, y tratar la restricción como la posición que es.

¿La afirmación es sobre explicar o sobre describir? Al menos una posición desarrollada sostiene que las leyes que mejor unifican no son las que mejor describen. Se acepte o no, las dos virtudes merecen distinguirse.

¿El argumento necesita que los inobservables sean reales? A veces sí y a veces no, y un argumento que funciona en ambos casos es más fuerte que uno que necesita la cuestión resuelta.

¿Hay una formulación memorable soportando carga? Si la hay, conviene comprobarla en la fuente y no en la cita, porque las frases más repetidas son las menos examinadas.

Un modelo es una descripción deliberadamente falsa con una tarea. Juzgarlo consiste en preguntar qué tarea, qué se dejó fuera, qué se distorsionó a propósito y a qué se renunció para hacerlo utilizable.

En qué punto estamos

El artículo termina en preguntas para llevarse a un modelo: qué toma por objetivo, qué distorsiones son deliberadas, a qué se renunció, si la afirmación versa sobre explicar o sobre describir, si el argumento necesita que sus inobservables sean reales, y si una formulación memorable está soportando carga. Una descripción falsa deliberada con un trabajo que hacer. Qué sabemos, y cómo, es lo que sigue.

Lo que sabemos y cómo lo sabemos

La formulación de referencia enuncia que muchos modelos científicos son representacionales y están en lugar de una parte o un aspecto seleccionados del mundo, y que los modelos idealizados implican una simplificación o distorsión deliberada orientada a hacerlos tratables. De Cartwright se recoge, desde la descripción editorial del libro y no desde su argumento, que las leyes explicativas fundamentales no describen regularidades de la naturaleza mientras que las leyes localizadas y las entidades teóricas sí pueden leerse en clave realista. Levins escribió en 1966 que generalidad, realismo y precisión no pueden maximizarse a la vez en un modelo manejable, y realismo significa allí detalle conservado. Orzack y Sober rechazan la necesidad de ese intercambio. El artículo de 1976

argumenta contra la elaboración y una comprobación directa confirma que la frase famosa no está en él; la redacción literal aparece en un informe técnico de 1979. La obra de referencia da el realismo científico como creencia en el contenido observable e inobservable, y la adecuación empírica de van Fraassen, citada dentro de esa misma entrada, como verdad sobre lo observable.

Lo que todavía no sabemos

Si las partes inobservables de un modelo que funciona corresponden a algo real, cuestión para la que este artículo enuncia los dos criterios y no zanja ninguno. Si la posición de Cartwright resiste una lectura del propio argumento, ya que lo que aquí se recoge procede de una descripción editorial. Con qué frecuencia se atribuye mal la frase famosa, para lo que no se consultó ningún estudio bibliométrico. Y si, en un caso concreto, generalidad, detalle conservado y precisión se intercambian de hecho, que es lo que la objeción deja abierto al negar solamente que deban hacerlo.

Qué podría hacernos cambiar de opinión

Una lectura del argumento de Cartwright, y no de la descripción de su editorial, zanjaría si la posición es la que aquí se enuncia. Un estudio bibliométrico convertiría la atribución errónea de observación en frecuencia medida. Sobre la restricción, un caso en que las tres propiedades mejoren a la vez no refutaría a Levins tal como él la formuló, pero demostrar que deben intercambiarse en alguna clase concreta de modelos estrecharía la objeción. Sobre el realismo, un argumento que muestre que el éxito autoriza la creencia sin suponer lo que quiere probar movería una cuestión que este artículo deja abierta.

Si solo recuerdas una cosa

Un modelo es una descripción falsa deliberada con un trabajo, así que pregunta qué trabajo, qué se dejó fuera, qué se distorsionó a propósito y qué costó.

Artículos relacionados

- KNO-006 · *Mapas y representación espacial*

Referencias

Los números entre corchetes del cuerpo son números de afirmación y no de referencia. Cada uno remite a una fila del registro de afirmaciones, que nombra la fuente de abajo que la sostiene.

Box, G. E. P. 1979. *Robustness in the Strategy of Scientific Model Building*. MRC Technical Summary Report #1954. Mathematics Research Center, University of Wisconsin–Madison, mayo de 1979. **Sin DOI ni ISBN: informe técnico**. URL estable del ejemplar digitalizado consultado: https://archive.org/details/DTIC_ADA070213

Box, George E. P. 1976. «Science and Statistics». *Journal of the American Statistical Association* 71 (356): 791–799. **Sin DOI localizado sobre página consultada**. URL estable del ejemplar consultado: <https://www-sop.inria.fr/members/Ian.Jermyn/philosophy/writings/Boxonmaths.pdf>

Cartwright, Nancy. 1983. *How the Laws of Physics Lie*. Oxford University Press. **Manifestación en línea**, ISBN 9780191597152. DOI 10.1093/0198247044.001.0001

Chakravartty, Anjan. «Scientific Realism». *Stanford Encyclopedia of Philosophy*. Publicada por primera vez el 27 de abril de 2011; revisión sustantiva del 12 de junio de 2017. URL estable de la versión archivada consultada: <https://plato.stanford.edu/archives/sum2017/entries/scientific-realism/>

Frigg, Roman, y Stephan Hartmann. «Models in Science». *Stanford Encyclopedia of Philosophy*. Publicada por primera vez el 27 de febrero de 2006; revisión sustantiva del 2 de abril de 2025. URL estable de la versión archivada consultada: <https://plato.stanford.edu/archives/sum2025/entries/models-science/>

Levins, Richard. 1966. «The Strategy of Model Building in Population Biology». *American Scientist* 54 (4): 421–431. **Sin DOI: la obra es anterior al sistema DOI y ninguna página consultada muestra uno asignado con posterioridad.** URL estable indicada por el propio ejemplar consultado: <https://www.jstor.org/stable/27836590>

Orzack, Steven Hecht, y Elliott Sober. 1993. «A Critical Assessment of Levins's "The Strategy of Model Building in Population Biology" (1966)». *The Quarterly Review of Biology* 68 (4): 533–546. DOI 10.1086/418301

KNO-014 · Causalidad y correlación

KNO · Formas de conocer

Juan Segurola

v0.7 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia al publicar:** Todos los derechos reservados

Resumen del artículo

Cuando dos variables cambian juntas, ¿qué evidencia adicional hace falta antes de sostener que modificar una cambiaría la otra? Este artículo examina esa pregunta mediante orden temporal, confusión, causalidad inversa, aleatorización, diagramas causales, ajuste, puntuaciones de propensión, variables instrumentales, controles negativos y emulación de ensayos objetivo. Pregunta cómo cambia cada método el problema inferencial, qué supuestos permanecen ocultos cuando una asociación se trata como autoexplicativa y por qué distintos diseños identifican contrastes causales distintos. También pregunta cuándo la evidencia observacional puede sustentar inferencia causal sin fingir que reproduce un ensayo aleatorizado y cuándo los métodos formales solo vuelven más explícitos los supuestos. El objetivo es separar el patrón observable de la pregunta de intervención y mostrar qué debe quedar inspeccionable en una afirmación causal defendible.

En 60 segundos

Si dos variables cambian conjuntamente, varias estructuras pueden explicar el patrón. Una puede causar la otra. Una tercera variable puede influir en ambas. El resultado puede causar parcialmente la supuesta exposición. La selección de quién entra en el conjunto de datos puede inducir una asociación. El azar y los procesos de medida también pueden contribuir. Por ello, la inferencia causal empieza definiendo una pregunta causal, la intervención o contraste de interés y qué diseño o supuestos permitirían que los datos observados respondieran a esa pregunta.

La aleatorización es potente porque la asignación al azar puede, si se implementa correctamente, romper los vínculos sistemáticos entre la asignación del tratamiento y las causas basales del resultado. Los estudios observacionales no disponen de esa protección por diseño, pero no están limitados a describir correlaciones. Pueden recurrir al diseño, al conocimiento sustantivo, a diagramas causales, emparejamiento o ponderación, variables instrumentales, controles negativos, análisis de sensibilidad y emulación explícita de un ensayo objetivo. Ninguna de estas técnicas convierte automáticamente los datos en causalidad. Cada una identifica un efecto causal solo bajo determinadas condiciones.

Por qué importa esta pregunta

La medicina, la ingeniería, la economía, la ciencia del clima y la política pública necesitan a menudo saber qué ocurriría si se cambiara una exposición, tratamiento, variable de proceso o política. Una asociación puede motivar esa pregunta, pero no responderla por sí sola. La diferencia importa porque

predecir e intervenir no son la misma tarea. Comprender el puente inferencial entre correlación y causalidad ayuda a ver por qué diseño, tiempo, supuestos causales y diagnósticos de sesgo forman parte de la evidencia.

1. La asociación es un patrón; la causalidad es una pregunta contrafactual

Una asociación estadística describe cómo covarían unas variables en los datos observados. Un efecto causal compara resultados bajo distintas condiciones posibles de exposición o tratamiento. La formulación epidemiológica de Hernán hace explícita esta diferencia: los efectos causales se definen mediante contrastes entre resultados potenciales, mientras que las asociaciones observadas pueden diferir de esos contrastes porque los grupos expuestos de forma distinta pueden diferir ya en aspectos relevantes [2]. El tratamiento de Holland del modelo causal de Rubin sitúa la misma separación en el centro del razonamiento estadístico sobre causalidad [3].

La formulación contrafactual también revela un obstáculo fundamental. Para una misma unidad y en un mismo momento no observamos simultáneamente el resultado bajo tratamiento y sin tratamiento. El contrafactual que falta no es una molestia técnica que desaparezca al aumentar el tamaño de la base de datos. La inferencia causal depende por tanto de comparaciones entre unidades, tiempos o regímenes junto con supuestos que hagan informativas esas comparaciones.

Afirmación KNO014-C01. Una asociación observada y un efecto causal son cantidades distintas; los efectos causales requieren un contraste entre resultados potenciales o regímenes de intervención, no solo una covarianza o un coeficiente de regresión. [2,3]

Afirmación KNO014-C02. El problema causal fundamental es que, para una misma unidad, solo se observa uno de los resultados potenciales relevantes. [3]

En qué punto estamos

Una asociación estadística describe cómo covarían unas variables en los datos observados. Para una misma unidad y en un mismo momento no observamos simultáneamente el resultado bajo tratamiento y sin tratamiento. La inferencia causal depende por tanto de comparaciones entre unidades, tiempos o regímenes junto con supuestos que hagan informativas esas comparaciones. La sección siguiente pregunta por qué el orden temporal ayuda pero no resuelve la cuestión.

2. Por qué el orden temporal ayuda pero no resuelve la cuestión

Una causa debe preceder a su efecto dentro de la relación causal estudiada. Esto convierte el orden temporal en una restricción esencial. Si una supuesta causa ocurre únicamente después del resultado, esa afirmación falla para esa relación temporal.

Pero la precedencia no basta. Una exposición puede preceder a un resultado y, aun así, ambas estar impulsadas por una causa común anterior. Un biomarcador puede cambiar antes del diagnóstico mientras ya refleja un proceso patológico oculto. Una política puede adoptarse antes de un cambio económico porque los gobiernos anticipan ese cambio. El orden temporal elimina algunas historias causales, no todas las historias competidoras.

Bradford Hill trató la temporalidad como una consideración al evaluar si una asociación observada podía ser causal, pero también advirtió contra la interpretación de sus consideraciones como criterios rígidos que demostraran causalidad mecánicamente [1]. Su marco sigue siendo históricamente importante precisamente porque abordaba la inferencia a partir de evidencia observacional imperfecta, no porque proporcionara una lista universal de comprobación.

Afirmación KNO014-C03. La causa debe preceder al efecto en el orden causal, pero el orden temporal por sí solo no es suficiente para identificar un efecto causal. [2]

En qué punto estamos

Una causa debe preceder a su efecto dentro de la relación causal estudiada. Pero la precedencia no basta. Si una supuesta causa ocurre únicamente después del resultado, esa afirmación falla para esa relación temporal. Esto convierte el orden temporal en una restricción esencial. El orden temporal elimina algunas historias causales, no todas las historias competidoras. La sección siguiente examina confusión.

3. Confusión: cuando una causa común produce la asociación

Supongamos que fumar está asociado con un resultado, pero la edad influye tanto en el historial de tabaquismo como en ese resultado. Si la edad no se trata adecuadamente, parte de la asociación observada puede reflejar la edad en lugar del efecto causal del tabaquismo.

Esto es confusión en su forma más simple: una causa común de la exposición y el resultado abre un camino no causal entre ambos. Los diagramas causales modernos hacen explícita esta estructura. Greenland, Pearl y Robins mostraron cómo los grafos acíclicos dirigidos pueden codificar relaciones causales supuestas y ayudar a identificar las variables que deben controlarse para estimar un efecto no confundido [4].

La palabra importante es «supuestas». Un grafo no se descubre simplemente ajustando los datos. Sus flechas representan compromisos causales aportados por conocimiento científico, información de diseño y restricciones temporales. Los datos pueden contradecir algunas consecuencias de un grafo, pero un grafo compatible con los datos no queda por ello demostrado como verdadero.

Afirmación KNO014-C04. La confusión puede producir una asociación exposición-resultado aunque parte o toda esa asociación no sea el efecto causal de la exposición. [4]

Afirmación KNO014-C05. Los diagramas causales pueden hacer explícitos los supuestos de ajuste, pero la estructura causal codificada por sus flechas requiere supuestos sustantivos adicionales a la distribución conjunta observada. [4,5]

En qué punto estamos

Supongamos que fumar está asociado con un resultado, pero la edad influye tanto en el historial de tabaquismo como en ese resultado. Si la edad no se trata adecuadamente, parte de la asociación observada puede reflejar la edad en lugar del efecto causal del tabaquismo. La palabra importante es «supuestas». Los diagramas causales modernos hacen explícita esta estructura. La sección siguiente examina causalidad inversa.

4. Causalidad inversa: cuando la flecha apunta en la otra dirección

Una asociación entre X e Y no nos dice si X causa Y, Y causa X, ocurren ambas direcciones o una tercera estructura explica la dependencia. La causalidad inversa es especialmente importante cuando los procesos patológicos comienzan antes del diagnóstico. La pérdida de peso puede predecir una enfermedad posterior porque una enfermedad preclínica produjo la pérdida de peso, no porque la pérdida de peso causara la enfermedad.

Los datos longitudinales ayudan porque establecen secuencias, pero no eliminan automáticamente la causalidad inversa. El proceso causal relevante puede haber empezado antes de la primera medición registrada. La respuesta correcta es definir con precisión el tiempo cero, usar conocimiento sobre latencia e historia de la enfermedad y comprobar si estructuras temporales plausibles modifican la inferencia.

Afirmación KNO014-C06. El orden longitudinal puede restringir explicaciones por causalidad inversa, pero no excluir por sí solo procesos causales iniciados antes de la medición. [2,7]

En qué punto estamos

Una asociación entre X e Y no nos dice si X causa Y, Y causa X, ocurren ambas direcciones o una tercera estructura explica la dependencia. La pérdida de peso puede predecir una enfermedad posterior porque una enfermedad preclínica produjo la pérdida de peso, no porque la pérdida de peso causara la enfermedad. La sección siguiente examina aleatorización.

5. Aleatorización: por qué un experimento cambia el problema probatorio

En un experimento aleatorizado ideal, la asignación del tratamiento se genera independientemente de las causas basales del resultado. Con tamaño adecuado, ocultación de la asignación, cumplimiento del protocolo y medida válida del resultado, la aleatorización puede hacer comparables los grupos en expectativa y permitir estimar efectos causales sin los mismos supuestos de ausencia de confusión exigidos por una comparación observacional ordinaria [2].

Esto no significa que todo ensayo aleatorizado esté libre de sesgo. Las pérdidas de seguimiento, el incumplimiento, la clasificación errónea del resultado, la selección posterior a la aleatorización y las decisiones analíticas pueden seguir generando problemas. Tampoco puede aleatorizarse cualquier pregunta causal: algunas intervenciones son poco éticas, imposibles, demasiado raras o demasiado prolongadas.

El punto epistémico es más concreto. La aleatorización cambia el mecanismo de asignación. No hace visible la causalidad; crea un diseño en el que un contraste causal específico puede identificarse bajo un conjunto más pequeño y transparente de supuestos.

Afirmación KNO014-C07. Una aleatorización correctamente implementada puede identificar efectos causales al romper causas basales sistemáticas de la asignación, pero la validez del ensayo sigue dependiendo de implementación, seguimiento, medida y análisis. [2]

En qué punto estamos

En un experimento aleatorizado ideal, la asignación del tratamiento se genera independientemente de las causas basales del resultado. La aleatorización reduce la dependencia sistemática entre tratamiento y causas basales, pero no elimina problemas de implementación, seguimiento, medida o análisis. La sección siguiente examina por qué ajustar más variables no equivale necesariamente a controlar mejor el sesgo.

6. Ajustar no significa «controlarlo todo»

Una intuición frecuente afirma que un estudio observacional se vuelve más causal cuanto más variables ajusta. Los diagramas causales muestran por qué esto es falso.

Si una variable es confusora, ajustarla puede bloquear un camino trasero no causal. Si está en la vía causal, ajustarla puede eliminar parte del efecto total que se pretendía estimar. Si es un colisionador, es decir, está causada por dos variables de un camino, condicionar sobre ella puede abrir una asociación que no existía antes. En esos casos, «controlar» una variable fabrica sesgo.

Greenland, Pearl y Robins formalizaron cómo la estructura gráfica guía el ajuste [4]. Trabajos metodológicos posteriores han subrayado que controlar colisionadores, mediadores o indiscriminadamente todas las variables previas al tratamiento puede distorsionar el efecto causal cuando la estructura es inadecuada [6].

Afirmación KNO014-C08. Más ajuste estadístico no es monótonamente mejor: ajustar un confusor puede reducir sesgo, mientras que ajustar un mediador o colisionador puede modificar o introducir sesgo según el estimando y la estructura causal. [4,6]

En qué punto estamos

Una intuición frecuente afirma que un estudio observacional se vuelve más causal cuanto más variables ajusta. Si una variable es confusora, ajustarla puede bloquear un camino trasero no causal. Si está en la vía causal, ajustarla puede eliminar parte del efecto total que se pretendía estimar. La sección siguiente examina la idea de que resultados potenciales y diagramas causales son lenguajes complementarios.

7. Resultados potenciales y diagramas causales son lenguajes complementarios

Dos grandes tradiciones modernas describen la inferencia causal mediante lenguajes distintos pero solapados. Los marcos de resultados potenciales definen efectos comparando resultados bajo estados alternativos de tratamiento. Los modelos causales estructurales y los DAG representan gráficamente y algebraicamente las relaciones causales y la estructura de intervención.

No son simplemente vocabularios rivales entre los que haya que elegir. En muchos análisis prácticos iluminan partes diferentes del mismo problema. Los resultados potenciales obligan a precisar el contraste causal y la población objetivo. Los DAG obligan a precisar las vías supuestas, la confusión, la mediación y la selección. Ambos hacen más visibles los supuestos que un coeficiente de regresión presentado sin modelo causal.

Pearl destacó que las afirmaciones causales requieren supuestos no contenidos en las distribuciones probabilísticas ordinarias y que hacen falta lenguajes causales para expresarlos [5]. Hernán y Robins organizan igualmente la inferencia causal alrededor de estrategias de tratamiento y resultados contrafactuales explícitamente especificados [8].

Afirmación KNO014-C09. Los enfoques de resultados potenciales y los enfoques gráficos ofrecen formas complementarias de expresar estimandos y supuestos causales; ninguno reduce la inferencia causal a asociación ordinaria. [5,8]

En qué punto estamos

Dos grandes tradiciones modernas describen la inferencia causal mediante lenguajes distintos pero solapados. No son simplemente vocabularios rivales entre los que haya que elegir. Los modelos causales estructurales y los DAG representan gráficamente y algebraicamente las relaciones causales y la estructura de intervención. Hernán y Robins organizan igualmente la inferencia causal alrededor de estrategias de tratamiento y resultados contrafactuales explícitamente especificados [8]. La sección siguiente examina propensity scores.

8. Propensity scores: equilibrar covariables medidas, no la realidad desconocida

Rosenbaum y Rubin introdujeron el propensity score como la probabilidad de asignación al tratamiento condicionada a covariables observadas [9]. Bajo sus supuestos, equilibrar por el propensity score puede equilibrar esas covariables observadas entre grupos y apoyar la estimación de efectos causales.

Pero un propensity score no resuelve automáticamente la confusión no medida. Si una causa común importante no está entre las covariables observadas, equilibrar las variables medidas no equilibra la causa ausente. El método es por tanto potente dentro de sus condiciones de identificación e impotente frente a violaciones que no puede observar.

Afirmación KNO014-C10. Los métodos de propensity score pueden equilibrar covariables observadas bajo determinados supuestos, pero no eliminan por sí solos el sesgo de confusión no medida. [9]

En qué punto estamos

Rosenbaum y Rubin introdujeron el propensity score como la probabilidad de asignación al tratamiento condicionada a covariables observadas [9]. Bajo sus supuestos, equilibrar por el propensity score puede equilibrar esas covariables observadas entre grupos y apoyar la estimación de efectos causales. El método es por tanto potente dentro de sus condiciones de identificación e impotente frente a violaciones que no puede observar. La sección siguiente examina variables instrumentales.

9. Variables instrumentales: otra ruta con otros supuestos

A veces se utiliza una variable instrumental: una variable que modifica el tratamiento o la exposición pero que, bajo el modelo de identificación, afecta al resultado solo a través de ese tratamiento y satisface además condiciones de independencia y relevancia.

Angrist, Imbens y Rubin mostraron cómo el razonamiento con variables instrumentales puede identificar un efecto medio local del tratamiento para los cumplidores bajo supuestos explícitos, incluida la monotonicidad en su formulación [10]. El resultado ilustra una lección recurrente: los métodos no producen «el efecto causal» en abstracto. Identifican estimandos concretos para poblaciones concretas bajo supuestos concretos.

Afirmación KNO014-C11. Los métodos de variables instrumentales pueden identificar efectos causales sin el supuesto ordinario de ausencia de confusión no medida, pero sustituyéndolo por otros supuestos fuertes de identificación y, a menudo, por un estimando más específico. [10]

En qué punto estamos

A veces se utiliza una variable instrumental: una variable que modifica el tratamiento o la exposición pero que, bajo el modelo de identificación, afecta al resultado solo a través de ese tratamiento y satisface además condiciones de independencia y relevancia. Identifican estimandos concretos para poblaciones concretas bajo supuestos concretos. La sección siguiente examina controles negativos.

10. Controles negativos: buscar sesgo donde no debería existir efecto

Una exposición o resultado de control negativo se elige de modo que, bajo la hipótesis causal evaluada, el efecto causal objetivo deba ser nulo. Si aun así aparece una asociación, puede revelar confusión residual, sesgo de selección u otro error sistemático.

Lipsitch, Tchetgen Tchetgen y Cohen propusieron controles negativos como diagnóstico sistemático para estudios causales observacionales [11]. Un control negativo no es un falsador universal. Su interpretación depende de por qué se espera que carezca de efecto causal y de si comparte el mecanismo de sesgo que se intenta detectar. Bien usado, convierte supuestos de fondo en una prueba empírica de estrés.

Afirmación KNO014-C12. Los controles negativos pueden detectar algunas formas de confusión residual o sesgo cuando el control es causalmente nulo para la relación objetivo pero comparte vías relevantes de sesgo. [11]

En qué punto estamos

Una exposición o resultado de control negativo se elige de modo que, bajo la hipótesis causal evaluada, el efecto causal objetivo deba ser nulo. Si aun así aparece una asociación, puede revelar confusión residual, sesgo de selección u otro error sistemático. Bien usado, convierte supuestos de fondo en una prueba empírica de estrés. La sección siguiente examina ensayos objetivo.

11. Ensayos objetivo: diseñar la pregunta observacional antes de analizar los datos

Muchos estudios observacionales intentan implícitamente responder a la pregunta que habría respondido un ensayo aleatorizado si ese ensayo fuese posible. Hernán y Robins propusieron hacer explícito ese ensayo hipotético: criterios de elegibilidad, estrategias de tratamiento, procedimiento de asignación, tiempo cero, seguimiento, resultado, contraste causal y plan de análisis se especifican primero; después, el análisis observacional intenta emularlos lo mejor posible [7].

Este enfoque puede revelar defectos de diseño evitables. Si la elegibilidad se evalúa en un momento, el tratamiento en otro y el seguimiento empieza en un tercero, el análisis puede introducir tiempo inmortal o sesgo de selección. La emulación de ensayo objetivo desplaza así la inferencia causal hacia aguas arriba: el diseño precede al modelado.

Afirmación KNO014-C13. La emulación explícita de un ensayo objetivo puede revelar desajustes temporales y de diseño en análisis observacionales antes de iniciar el modelado estadístico. [7,12]

En qué punto estamos

Muchos estudios observacionales intentan implícitamente responder a la pregunta que habría respondido un ensayo aleatorizado si ese ensayo fuese posible. Este enfoque puede revelar defectos de diseño evitables. Si la elegibilidad se evalúa en un momento, el tratamiento en otro y el seguimiento empieza en un tercero, el análisis puede introducir tiempo inmortal o sesgo de selección. La sección siguiente revisa el marco de integración de evidencia de Bradford Hill.

12. Bradford Hill revisitado: integrar evidencia, no puntuar causalidad

La discusión de Hill de 1965 sigue siendo influyente porque trató cómo razonar cuando no hay experimentos controlados disponibles. Fuerza, consistencia, especificidad, temporalidad, gradiente biológico, plausibilidad, coherencia, experimento y analogía se ofrecieron como puntos de vista a considerar, no como una puntuación aditiva validada ni como condiciones necesarias y suficientes [1].

La inferencia causal moderna incorpora herramientas que Hill no tenía: estimandos contrafactuales, DAG, g-métodos, propensity scores, identificación por variables instrumentales, controles negativos y emulación de ensayos objetivo. Pero su lección general sigue siendo compatible con ellas: el juicio causal integra múltiples tipos de evidencia y evita inferir automáticamente causalidad a partir de una asociación.

La actualización importante es metodológica. Una asociación fuerte puede seguir estando confundida. Una asociación débil puede representar un efecto causal importante. La plausibilidad biológica cambia con el conocimiento de fondo. La consistencia entre estudios puede reflejar un sesgo compartido. El valor probatorio de cada característica depende de la pregunta causal y del diseño que generó la evidencia.

Afirmación KNO014-C14. Las consideraciones de Bradford Hill son ayudas para el razonamiento causal, no una lista mecánica que demuestre o descarte causalidad. [1]

En qué punto estamos

La discusión de Hill de 1965 sigue siendo influyente porque trató cómo razonar cuando no hay experimentos controlados disponibles. Fuerza, consistencia, especificidad, temporalidad, gradiente biológico, plausibilidad, coherencia, experimento y analogía se ofrecieron como puntos de vista a considerar, no como una puntuación aditiva validada ni como condiciones necesarias y suficientes [1]. La sección siguiente pregunta qué no pueden decir los datos por sí solos.

13. Lo que los datos por sí solos no pueden decir

Una base de datos suficientemente rica puede estimar muchas asociaciones condicionales con gran precisión. Esa precisión no identifica la estructura causal que las generó. Modelos causales distintos pueden implicar la misma distribución observada. La diferencia entre ellos aparece cuando preguntamos por intervenciones o contrafactuales.

Por eso la inferencia causal no puede reducirse a «usar un modelo más potente». El aprendizaje automático puede estimar funciones predictivas con enorme precisión y dejar sin identificar el estimando causal. A la inversa, un análisis sencillo puede sostener una afirmación causal fuerte cuando el diseño y los supuestos hacen identificable el contraste.

Afirmación KNO014-C15. Precisión predictiva e identificación causal son objetivos distintos; un modelo puede predecir bien sin establecer el efecto de una intervención. [5]

En qué punto estamos

Una base de datos suficientemente rica puede estimar muchas asociaciones condicionales con gran precisión. Esa precisión no identifica la estructura causal que las generó. Modelos causales distintos pueden implicar la misma distribución observada. La diferencia entre ellos aparece cuando preguntamos por intervenciones o contrafactuales. La sección siguiente convierte estos puntos en una secuencia práctica para el análisis.

14. Una secuencia práctica de inferencia causal

Un análisis disciplinado puede organizarse como una secuencia en lugar de un único test estadístico:

1. Formular la pregunta causal y la población objetivo.
2. Definir la exposición o intervención y el régimen de comparación.
3. Definir el resultado y el horizonte temporal.
4. Especificar el tiempo cero y el orden de los acontecimientos relevantes.
5. Dibujar o declarar la estructura causal supuesta.
6. Decidir qué variables son confusoras, mediadoras, colisionadores o consecuencias de la selección para el estimando concreto.
7. Elegir un diseño y un estimador cuyas condiciones de identificación correspondan a esa estructura.
8. Comprobar equilibrio, solapamiento, calidad de medida y datos faltantes.
9. Usar controles negativos, especificaciones alternativas o análisis de sensibilidad cuando proceda.
10. Interpretar la estimación como causal solo en la medida en que sigan siendo defendibles los supuestos de identificación y medida.

Esta secuencia no es una garantía. Es una forma de hacer inspeccionables las premisas ocultas de una afirmación causal.

Afirmación KNO014-C16. El análisis causal debe tratarse primero como un problema de diseño y supuestos y solo después como un problema de estimación. [4,7,8]

En qué punto estamos

Paso a paso, un análisis disciplinado define pregunta, comparación, resultado y tiempo cero; declara la estructura causal supuesta y el papel de cada variable; elige diseño y estimador acordes; comprueba equilibrio, solapamiento y medida; contrasta la estimación; y la interpreta como causal solo en la medida en que sus supuestos sigan siendo defendibles. No garantiza nada, pero hace inspeccionables las premisas ocultas. La sección siguiente delimita el alcance de KNO-014.

15. Frontera de KNO-014

KNO-014 trata el **problema metodológico de distinguir efectos causales de asociaciones no causales**. No pretende resolver la metafísica de la causalidad. Esa cuestión conceptual pertenece a IDE-004 · Causalidad.

Tampoco sustituye KNO-012 · Experimentos y control de variables. KNO-012 pregunta cómo crean los experimentos comparaciones controladas; KNO-014 pregunta cómo se identifican afirmaciones causales en contextos experimentales y observacionales.

KNO-015 · Incertidumbre y probabilidad abordará otro problema: cómo se representa y propaga la incertidumbre una vez especificados la cantidad objetivo y la estructura inferencial. La variabilidad de muestreo como explicación competidora de una asociación observada también se entrega a KNO-015 y no se desarrolla aquí.

En qué punto estamos

KNO-014 trata el problema metodológico de distinguir efectos causales de asociaciones no causales. No pretende resolver la metafísica de la causalidad. Esa cuestión conceptual pertenece a IDE-004 · Causalidad. Esto delimita el alcance del artículo y prepara la síntesis final sin ampliar la afirmación más allá de lo establecido en las secciones anteriores.

Lo que sabemos y cómo lo sabemos

La asociación y el efecto causal son objetivos distintos. La aleatorización, los diagramas causales, los resultados potenciales, los controles negativos, las variables instrumentales y la emulación de ensayos objetivo hacen inspeccionables partes diferentes del problema causal. Su fuerza probatoria procede del diseño y de los supuestos, no solo de la estimación estadística.

Lo que todavía no sabemos

Los datos observados no pueden establecer por sí solos la dirección causal, la ausencia de confusión no medida, una restricción de exclusión válida o la corrección de un grafo causal supuesto. Esas cuestiones requieren conocimiento externo, información de diseño y pruebas cuyo alcance también es limitado.

Qué podría hacernos cambiar de opinión

La confianza en una afirmación causal debe cambiar si nueva evidencia revela confusión oculta, orden temporal violado, sesgo de medida, controles negativos fallidos, falta de solapamiento, un instrumento inválido o una especificación de ensayo objetivo que ya no corresponde al proceso real.

Si solo recuerdas una cosa

La inferencia causal empieza cuando una asociación se conecta con una pregunta de intervención bien definida mediante diseño y supuestos explícitos y contrastables.

Artículos relacionados

- KNO-012 · Experiments and Variable Control
- KNO-013 · Models, Maps and Reality
- KNO-015 · Uncertainty and Probability
- KNO-016 · How Evidence Is Built: Putting the Tools Together

Referencias

- [1] Hill, Austin Bradford. "The Environment and Disease: Association or Causation?" *Proceedings of the Royal Society of Medicine* 58(5), 1965, 295–300. DOI: 10.1177/003591576505800503. PMID: 14283879.
- [2] Hernán, M. A. "A Definition of Causal Effect for Epidemiological Research." *Journal of Epidemiology and Community Health* 58(4), 2004, 265–271. DOI: 10.1136/jech.2002.006361. PMID: 15026432.
- [3] Holland, Paul W. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396), 1986, 945–960. DOI: 10.1080/01621459.1986.10478354.
- [4] Greenland, Sander; Pearl, Judea; Robins, James M. "Causal Diagrams for Epidemiologic Research." *Epidemiology* 10(1), 1999, 37–48. DOI: 10.1097/00001648-199901000-00008. PMID: 9888278.
- [5] Pearl, Judea. "An Introduction to Causal Inference." *The International Journal of Biostatistics* 6(2), 2010, Article 7. DOI: 10.2202/1557-4679.1203. PMID: 20305706.
- [6] Tennant, Peter W. G.; Murray, Eleanor J.; Arnold, Kellyn F.; Berrie, Laurie; Fox, Matthew P.; Gadd, Sarah C.; Harrison, Wendy J.; Keeble, Claire; Ranker, Lynsie R.; Textor, Johannes; Tomova, Georgia D.; Gilthorpe, Mark S.; Ellison, George T. H. "Use of Directed Acyclic Graphs (DAGs) to Identify Confounders in Applied Health Research: Review and Recommendations." *International Journal of Epidemiology* 50(2), 2021, 620–632. DOI: 10.1093/ije/dyaa213.
- [7] Hernán, Miguel A.; Robins, James M. "Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available." *American Journal of Epidemiology* 183(8), 2016, 758–764. DOI: 10.1093/aje/kwv254. PMID: 26994063.
- [8] Hernán, Miguel A.; Robins, James M. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC, 2020. Cita sugerida por los autores; el texto en línea se consultó por separado como manifestación electrónica revisable.
- [9] Rosenbaum, Paul R.; Rubin, Donald B. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1), 1983, 41–55. DOI: 10.1093/biomet/70.1.41.
- [10] Angrist, Joshua D.; Imbens, Guido W.; Rubin, Donald B. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434), 1996, 444–455. DOI: 10.1080/01621459.1996.10476902.

[11] Lipsitch, Marc; Tchetgen Tchetgen, Eric; Cohen, Ted. "Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies." *Epidemiology* 21(3), 2010, 383–388. DOI: 10.1097/EDE.0b013e3181d61eeb. PMID: 20335814.

[12] Hernán, Miguel A.; Sauer, Brian C.; Hernández-Díaz, Sonia; Platt, Robert; Shrier, Ian. "Specifying a Target Trial Prevents Immortal Time Bias and Other Self-Inflicted Injuries in Observational Analyses." *Journal of Clinical Epidemiology* 79, 2016, 70–75. DOI: 10.1016/j.jclinepi.2016.04.014. PMID: 27237061.

KNO-015 · Incertidumbre y probabilidad

KNO · Formas de conocer

Juan Segurola

v0.7 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia al publicar:** Todos los derechos reservados

Resumen del artículo

Los resultados científicos suelen venir acompañados de intervalos, probabilidades, barras de error o análisis de sensibilidad, pero ¿qué clases de incertidumbre representan realmente esas herramientas? Este artículo pregunta en qué se diferencian la incertidumbre de medida, la variación muestral, la incertidumbre paramétrica, la incertidumbre de modelo, la incertidumbre estructural y el juicio experto, y qué ocurre cuando se comprimen en un solo número. Examina intervalos de confianza, intervalos creíbles, valores p, propagación, predicción y análisis de sensibilidad preguntando qué afirmaciones dependen del marco estadístico elegido y qué límites permanecen fuera de él. También pregunta cómo comunicar incertidumbre sin convertir ignorancia en falsa precisión ni tratar todo desconocimiento como cuantificable. El problema central es metodológico: qué representación de la incertidumbre corresponde a la afirmación y qué queda sin representar aun después del cálculo.

En 60 segundos

Un número sin declaración de incertidumbre puede parecer más exacto de lo que permite la evidencia. Sin embargo, añadir «±» no basta. El intervalo puede describir el comportamiento en muestreos repetidos, una distribución de probabilidad asignada a un mensurando, una distribución posterior bayesiana, una distribución predictiva o una convención basada en un factor de cobertura. Son objetos que responden a preguntas distintas.

La disciplina práctica es, por tanto: identificar la magnitud de interés; enumerar las fuentes materiales de incertidumbre; decidir cuáles pueden representarse probabilísticamente; especificar el marco inferencial; propagar la incertidumbre a través del modelo de medida o estadístico; examinar la sensibilidad a los supuestos; y explicar qué significa realmente el intervalo o la probabilidad resultante. La probabilidad puede organizar la incertidumbre. No puede borrar los supuestos, las omisiones desconocidas ni la mala especificación del modelo. La idea práctica es separar el resultado observado de los supuestos que lo conectan con la interpretación. Cada herramienta reduce algunas ambigüedades y deja otras abiertas; por eso el análisis debe declarar qué puede identificar, qué depende de condiciones adicionales y qué observación podría contradecir la conclusión. La precisión formal no elimina por sí sola esos límites.

Por qué importa esta pregunta

Las decisiones científicas suelen tomarse bajo incertidumbre, pero incertidumbres distintas exigen respuestas distintas. Un intervalo estrecho puede coexistir con un modelo mal especificado; un instrumento preciso puede medir la cantidad equivocada; una probabilidad puede estar bien calculada y excluir el desconocimiento más importante. Tratar todas las incertidumbres como intercambiables produce falsa precisión. Distinguir qué se mide, qué se muestrea, qué se modeliza y qué queda fuera del modelo es esencial para interpretar la evidencia y comunicar la fuerza de una conclusión.

1. La incertidumbre no es una sola sustancia

La palabra cubre mecanismos distintos que no deben colapsarse en una categoría residual única.

[C01] La incertidumbre de medida se refiere al conocimiento incompleto del valor de un mensurando dados el procedimiento de medida, las entradas, las correcciones y el modelo. El marco GUM establece reglas generales para evaluar y expresar esa incertidumbre, y documentos posteriores del JCGM amplían el marco a la propagación de distribuciones y a los modelos de medida. [S01–S03]

[C02] La incertidumbre muestral aparece porque una muestra finita es solo una realización de un proceso generador de datos más amplio. Su tratamiento estadístico depende del mecanismo de muestreo o asignación y del estimador estudiado. Es conceptualmente distinta de la incertidumbre de calibración de un instrumento o de la incertidumbre estructural del modelo. [S01, S08]

[C03] La incertidumbre del modelo incluye la forma funcional, la estructura omitida, las aproximaciones y las especificaciones alternativas plausibles. La incertidumbre estructural no siempre puede reducirse a un único error estándar calculado dentro de un modelo elegido. [S07, S09]

Estas categorías pueden interactuar. Una constante de calibración puede entrar en un modelo estadístico; un modelo puede transformar entradas medidas; un diseño muestral puede determinar qué parámetros son estimables. La exigencia metodológica no consiste en aislar artificialmente cada fuente, sino en evitar que un componente cuantificado se confunda con toda la incertidumbre.

En qué punto estamos

La palabra cubre mecanismos distintos que no deben colapsarse en una categoría residual única. Las incertidumbres de medida, muestral y del modelo son conceptualmente distintas, pero pueden interactuar, como cuando una constante de calibración entra en un modelo estadístico. Lo decisivo es que ningún componente cuantificado se confunda con toda la incertidumbre. La sección siguiente examina la idea de que un resultado de medida no son solo sus dígitos mostrados.

2. Un resultado de medida no son solo sus dígitos mostrados

Un instrumento digital puede mostrar 12,345678 sin establecer conocimiento hasta seis decimales.

[C04] En el enfoque GUM/NIST, los componentes individuales de incertidumbre estándar se combinan, incluida la covarianza cuando corresponde, para obtener una incertidumbre estándar combinada. Cuando procede esa forma de comunicación, la incertidumbre expandida se obtiene multiplicando la incertidumbre estándar combinada por un factor de cobertura. [S01, S04]

[C05] Un factor de cobertura como $k = 2$ no garantiza mecánicamente una cobertura del 95 % en todos los casos. La interpretación probabilística depende de supuestos distribucionales y de otro tipo, y un número pequeño de grados de libertad efectivos puede exigir factores diferentes. [S04, S05]

La lección es general. La precisión que imprime un programa es formato. La precisión evidencial procede del modelo de incertidumbre y de la justificación de sus supuestos.

En qué punto estamos

Un instrumento digital puede mostrar 12,345678 sin establecer conocimiento hasta seis decimales. La precisión que imprime un programa es formato. La precisión evidencial procede del modelo de incertidumbre y de la justificación de sus supuestos. La sección siguiente examina la idea de que la probabilidad es una herramienta de representación, no un disolvente de la ignorancia.

3. La probabilidad es una herramienta de representación, no un disolvente de la ignorancia

La probabilidad entra en la ciencia con más de una función. En modelos de muestreo repetido caracteriza el comportamiento de variables aleatorias bajo repeticiones. En análisis bayesiano también se utiliza para representar incertidumbre sobre cantidades desconocidas condicionada a un modelo, una información previa y unos datos observados. En metrología, se emplean distribuciones de probabilidad para representar valores razonablemente atribuibles a cantidades inciertas y propagar esas distribuciones a través de modelos de medida. [S02, S03]

[C06] La misma probabilidad numérica puede, por tanto, proceder de construcciones inferenciales distintas. Un informe científicamente adecuado debe indicar a qué objeto aleatorio o cantidad incierta se refiere y bajo qué supuestos. [S02, S06]

KNO-015 no decide el significado filosófico de la probabilidad. Eso corresponde a IDE-008. Su preocupación es operacional: no pasar de una interpretación a otra sin declararlo.

En qué punto estamos

La probabilidad entra en la ciencia con más de una función. [S02, S03] KNO-015 no decide el significado filosófico de la probabilidad. En análisis bayesiano también se utiliza para representar incertidumbre sobre cantidades desconocidas condicionada a un modelo, una información previa y unos datos observados. La sección siguiente examina la idea de que los intervalos de confianza y los intervalos creíbles responden a preguntas distintas.

4. Los intervalos de confianza y los intervalos creíbles responden a preguntas distintas

La expresión «intervalo del 95 %» está incompleta.

[C07] Un procedimiento frecuentista de confianza se calibra mediante sus propiedades en muestreos repetidos bajo el modelo supuesto. La interpretación habitual del 95 % de confianza se refiere a la cobertura a largo plazo del procedimiento, no a una probabilidad posterior de que el parámetro fijo se encuentre en el intervalo concreto observado. [S10]

[C08] Un intervalo creíble bayesiano es una declaración probabilística sobre un parámetro o cantidad condicionada al modelo bayesiano especificado, incluidas la distribución previa y la verosimilitud. Por tanto, su interpretación no es intercambiable con la de un intervalo de confianza frecuentista, aunque sus extremos numéricos sean similares. [S10]

Ninguna etiqueta hace desaparecer los supuestos. La pregunta relevante es qué datos, modelo y construcción inferencial justifican el intervalo.

En qué punto estamos

La expresión «intervalo del 95 %» está incompleta. Ninguna etiqueta hace desaparecer los supuestos. La pregunta relevante es qué datos, modelo y construcción inferencial justifican el intervalo. Lo importante aquí es la relación declarada entre evidencia, método e inferencia, no una promesa genérica de certeza. La sección siguiente examina la idea de que un valor p no es la probabilidad de que una hipótesis sea verdadera.

5. Un valor p no es la probabilidad de que una hipótesis sea verdadera

Este error es tan frecuente que merece una frontera propia.

[C09] La American Statistical Association señala que los valores p pueden indicar incompatibilidad entre los datos y un modelo estadístico especificado, pero no miden la probabilidad de que la hipótesis estudiada sea verdadera, no miden el tamaño ni la importancia científica del efecto y no deberían determinar por sí solos conclusiones científicas o decisiones de política. [S06]

Un valor p se calcula bajo un modelo especificado, normalmente incluyendo una hipótesis nula y otros supuestos. No es una probabilidad posterior de esa hipótesis salvo que se construya adicionalmente un modelo bayesiano; en tal caso, la probabilidad posterior resultante es otro objeto.

En qué punto estamos

Este error es tan frecuente que merece una frontera propia. No es una probabilidad posterior de esa hipótesis salvo que se construya adicionalmente un modelo bayesiano; en tal caso, la probabilidad posterior resultante es otro objeto. La sección siguiente examina la idea de que la incertidumbre se propaga a través de modelos.

6. La incertidumbre se propaga a través de modelos

Si una salida depende de entradas inciertas, la incertidumbre de esas entradas se propaga hacia la salida.

[C10] La GUM proporciona reglas analíticas de propagación bajo sus condiciones de aplicabilidad, mientras JCGM 101 especifica la propagación de distribuciones de probabilidad a través de un modelo de medida mediante métodos de Monte Carlo. [S01, S02]

La propagación Monte Carlo no es magia. Hereda las distribuciones de entrada, la estructura de dependencia y el modelo de medida suministrados. Si son deficientes, una simulación enorme puede producir una respuesta muy precisa para un modelo de incertidumbre equivocado.

En qué punto estamos

Si una salida depende de entradas inciertas, la incertidumbre de esas entradas se propaga hacia la salida. La propagación Monte Carlo no es magia. Hereda las distribuciones de entrada, la estructura de dependencia y el modelo de medida suministrados. La elección del modelo determina también qué incertidumbre se transmite. La sección siguiente examina la idea de que la incertidumbre estructural necesita visibilidad propia.

7. La incertidumbre estructural necesita visibilidad propia

Algunas incertidumbres quedan fuera de la matriz de covarianza de parámetros del modelo seleccionado: mecanismos alternativos, condiciones de contorno, retroalimentaciones no representadas, dependencia desconocida o desacuerdo experto no resuelto.

[C11] La guía de incertidumbre del IPCC separa explícitamente la evaluación de evidencia y acuerdo, confianza, verosimilitud o probabilidad, y pide declarar supuestos y evaluar las incertidumbres estructurales cuando proceda. [S07]

Esta separación evita un error de categoría. Una distribución probabilística puede resumir incertidumbre condicionada a un modelo mientras la incertidumbre sobre el propio modelo permanece solo parcialmente cuantificada.

En qué punto estamos

Algunas incertidumbres quedan fuera de la matriz de covarianza de parámetros del modelo seleccionado: mecanismos alternativos, condiciones de contorno, retroalimentaciones no representadas, dependencia desconocida o desacuerdo experto no resuelto. Una distribución probabilística puede resumir incertidumbre condicionada a un modelo mientras la incertidumbre sobre el propio modelo permanece solo parcialmente cuantificada. La sección siguiente examina la idea de que el juicio experto puede disciplinarse sin fingir que es un dato observado.

8. El juicio experto puede disciplinarse sin fingir que es un dato observado

En algunos dominios hay pocas frecuencias directas, los modelos discrepan o las condiciones futuras no son repeticiones de un proceso histórico estable. El juicio experto puede entonces contribuir a la evaluación de incertidumbre.

[C12] La guía del IPCC utiliza lenguaje calibrado y una evaluación trazable de evidencia y acuerdo para disciplinar esos juicios, en vez de presentar toda declaración de confianza como si fuera una frecuencia observada en datos. [S07]

El juicio experto no es automáticamente más débil ni más fuerte que una salida estadística. Su función evidencial depende de transparencia, calibración, independencia, diseño de la elicitación y evidencia utilizada por los expertos. La regla crítica es la procedencia: el lector debe saber cuándo un número procede de observaciones, un conjunto de modelos, una elicitación formal o una combinación.

En qué punto estamos

En algunos dominios hay pocas frecuencias directas, los modelos discrepan o las condiciones futuras no son repeticiones de un proceso histórico estable. El juicio experto puede entonces contribuir a la evaluación de incertidumbre. La sección siguiente examina la idea de que los desconocidos desconocidos no están dentro de todos los intervalos.

9. Los desconocidos desconocidos no están dentro de todos los intervalos

Un presupuesto de incertidumbre solo puede incluir fuentes representadas en el análisis.

[C13] Los informes científicos deben distinguir, por tanto, la incertidumbre cuantificada de las limitaciones reconocidas pero no cuantificadas y del error genuinamente no reconocido. Las National Academies recomiendan comunicar incertidumbres reconocidas y fuentes potenciales en dimensiones estocásticas, de medida, computacionales, de conocimiento, de modelización y de análisis, en lugar de tratar un único intervalo estadístico como exhaustivo. [S08]

Eso no vuelve inútiles los intervalos cuantificados. Define su jurisdicción.

En qué punto estamos

Un presupuesto de incertidumbre solo puede incluir fuentes representadas en el análisis. Eso no vuelve inútiles los intervalos cuantificados. Su valor está en hacer inspeccionable este paso inferencial concreto, no en tratar el método como autojustificado. La sección siguiente examina la idea de que el análisis de sensibilidad prueba cuánto dependen las conclusiones de los supuestos.

10. El análisis de sensibilidad prueba cuánto dependen las conclusiones de los supuestos

Si una conclusión cambia drásticamente bajo especificaciones alternativas plausibles, la incertidumbre no queda capturada citando solo el error estándar de una especificación.

Un análisis de sensibilidad puede variar distribuciones previas, formas funcionales, parámetros auxiliares, supuestos sobre datos ausentes, entradas de calibración o elecciones estructurales plausibles. No es una cura universal: el analista todavía debe decidir qué alternativas son plausibles. Pero muestra dónde dependen las conclusiones de elecciones que de otro modo permanecerían ocultas.

En qué punto estamos

Si una conclusión cambia drásticamente bajo especificaciones alternativas plausibles, la incertidumbre no queda capturada citando solo el error estándar de una especificación. Un análisis de sensibilidad puede variar distribuciones previas, formas funcionales, parámetros auxiliares, supuestos sobre datos ausentes, entradas de calibración o elecciones estructurales plausibles. La sección siguiente examina la idea de que la incertidumbre predictiva no es la incertidumbre de parámetros.

11. La incertidumbre predictiva no es la incertidumbre de parámetros

Estimar una respuesta media y predecir una observación nueva son tareas distintas. Un parámetro del modelo puede estar bien estimado mientras los resultados futuros individuales siguen siendo muy variables.

[C14] Una declaración completa de incertidumbre debe identificar la cantidad objetivo: un parámetro, un estado latente, un mensurando, una observación futura, una media poblacional o un resultado relevante para una decisión. Objetivos distintos requieren en general construcciones de incertidumbre distintas. [S01, S03, S10]

Por eso «la incertidumbre es $\pm x$ » suele estar insuficientemente especificado.

En qué punto estamos

Estimar una respuesta media y predecir una observación nueva son tareas distintas. Un parámetro del modelo puede estar bien estimado mientras los resultados futuros individuales siguen siendo muy variables. Por eso «la incertidumbre es $\pm x$ » suele estar insuficientemente especificado. Esa diferencia cambia qué intervalo debe interpretarse. La sección siguiente examina un flujo de trabajo práctico para la incertidumbre.

12. Un flujo de trabajo práctico para la incertidumbre

Para el trabajo científico, la siguiente secuencia es más robusta que añadir un intervalo al final:

1. Definir la cantidad objetivo y el estimando o mensurando.
2. Identificar el proceso generador de datos o de medida.
3. Enumerar las fuentes materiales de incertidumbre y sus dependencias.
4. Separar variabilidad estocástica y conocimiento incompleto cuando la distinción sea relevante.
5. Especificar el modelo y el marco inferencial.
6. Cuantificar los componentes que puedan justificarse cuantitativamente.
7. Propagarlos a través del modelo, analítica o computacionalmente según proceda.
8. Probar la sensibilidad frente a supuestos alternativos plausibles.
9. Registrar limitaciones importantes que permanezcan sin cuantificar.
10. Explicar el significado de cada intervalo, probabilidad y declaración de cobertura.

El flujo no garantiza verdad. Hace auditable el camino desde la evidencia hasta la afirmación de incertidumbre.

En qué punto estamos

En lugar de añadir un intervalo al final, el flujo define la cantidad objetivo y su proceso de datos o medida, enumera fuentes de incertidumbre, separa variabilidad y conocimiento incompleto, especifica modelo y marco, cuantifica y propaga componentes justificados, prueba la sensibilidad, registra lo no cuantificado y explica qué significa cada intervalo. Hace auditable el camino sin garantizar verdad. La sección siguiente examina lo que la incertidumbre no autoriza.

13. Lo que la incertidumbre no autoriza

La incertidumbre no permite afirmar que todas las hipótesis sean igualmente plausibles. Tampoco un intervalo estrecho demuestra que un modelo sea correcto. Un resultado puede ser estadísticamente preciso y estar sistemáticamente equivocado. Un intervalo ancho puede reflejar honestamente información débil en vez de un fallo metodológico.

La mejor declaración de incertidumbre no es necesariamente el número más pequeño. Es aquella cuyo alcance, supuestos y construcción resisten la inspección.

En qué punto estamos

La incertidumbre no permite afirmar que todas las hipótesis sean igualmente plausibles. Un resultado puede ser estadísticamente preciso y estar sistemáticamente equivocado. Un intervalo ancho puede reflejar honestamente información débil en vez de un fallo metodológico. Esto delimita el alcance del artículo y prepara la síntesis final sin ampliar la afirmación más allá de lo establecido en las secciones anteriores.

Nota de frontera

KNO-015 trata **los métodos mediante los cuales el trabajo científico identifica, representa, propaga y comunica incertidumbre**. No pretende resolver qué es fundamentalmente la probabilidad; esa cuestión conceptual corresponde a **IDE-008 · Probability / Probabilidad**. Tampoco trata la incertidumbre como concepto ontológico o epistemológico general; eso corresponde a **IDE-017 · Uncertainty / Incertidumbre**. Se apoya directamente en **KNO-011 · Measuring Before Theorising**, **KNO-012 · Experiments and Variable Control**, **KNO-013 · Models, Maps and Reality** y **KNO-014 · Causality and Correlation**.

Enunciados de encuadre declarados no representados como afirmaciones numeradas

Esta lista es una ayuda de control, **no una garantía de exhaustividad semántica**. Registra enunciados sustantivos de encuadre que se han dejado deliberadamente fuera del registro numerado de afirmaciones y explica el motivo.

- «La incertidumbre no permite afirmar que todas las hipótesis sean igualmente plausibles». Encuadre normativo de la incertidumbre calibrada, no una afirmación empírica distinta.
- «Un resultado puede ser estadísticamente preciso y estar sistemáticamente equivocado». Advertencia metodológica condensada, ya delimitada por las secciones sobre incertidumbre sistemática y de modelo.
- «Un intervalo ancho puede reflejar honestamente información débil en vez de un fallo metodológico». Encuadre interpretativo de la amplitud del intervalo, no una regla cuantitativa independiente.
- «La mejor declaración de incertidumbre no es necesariamente el número más pequeño». Síntesis de las distinciones de incertidumbre anteriores, no una nueva proposición atribuida a una fuente.

Lo que sabemos y cómo lo sabemos

La incertidumbre puede proceder de la medida, el muestreo, los parámetros, la predicción, la estructura del modelo y el juicio, y esas fuentes no son intercambiables. Los intervalos y probabilidades estadísticas adquieren significado dentro de su marco inferencial, mientras el análisis de sensibilidad expone la dependencia de los supuestos.

Lo que todavía no sabemos

Ningún intervalo finito contiene automáticamente todo desconocimiento relevante. Omisiones estructurales, mecanismos no modelizados y posibilidades realmente imprevistas pueden permanecer fuera de un modelo probabilístico aunque la aritmética dentro de él sea correcta.

Qué podría hacernos cambiar de opinión

Una conclusión debe revisarse si cambia la calibración, nuevos datos amplían o reducen la incertidumbre muestral, estructuras de modelo alternativas ajustan de forma comparable, los análisis de sensibilidad revelan supuestos frágiles o mecanismos ignorados pasan a ser plausibles y relevantes.

Si solo recuerdas una cosa

Representa la incertidumbre en la forma que corresponde a la afirmación y declara qué desconocimientos importantes quedan fuera del cálculo.

Artículos relacionados

- KNO-011 · Measuring Before Theorising
- KNO-014 · Causality and Correlation
- KNO-016 · How Evidence Is Built: Putting the Tools Together
- KNO-021 · Open Science and Reproducibility

Referencias

[S01] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. JCGM 100:2008(E). DOI: 10.59161/JCGM100-2008E.

[S02] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method*. JCGM 101:2008. DOI: 10.59161/JCGM101-2008.

[S03] Joint Committee for Guides in Metrology (JCGM). *Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models*. JCGM GUM-6:2020. DOI: 10.59161/JCGMGUM-6-2020.

- [S04] Taylor, B. N.; Kuyatt, C. E. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. NIST Technical Note 1297. Manifestación web de NIST consultada el 5 de septiembre de 2026; página actualizada el 6 de mayo de 2026. <https://www.nist.gov/pml/nist-technical-note-1297>
- [S05] National Institute of Standards and Technology. *NIST TN 1297: Appendix B. Coverage Factors*. Página web de NIST, actualizada el 11 de septiembre de 2025; consultada el 5 de septiembre de 2026. <https://www.nist.gov/pml/nist-technical-note-1297/nist-tn-1297-appendix-b-coverage-factors>
- [S06] Wasserstein, Ronald L.; Lazar, Nicole A. “The ASA’s Statement on p-Values: Context, Process, and Purpose.” *The American Statistician* 70(2), 2016, 129–133. DOI: 10.1080/00031305.2016.1154108.
- [S07] Mastrandrea, Michael D.; Field, Christopher B.; Stocker, Thomas F.; Edenhofer, Ottmar; Ebi, Kristie L.; Frame, David J.; Held, Hermann; Kriegler, Elmar; Mach, Katharine J.; Matschoss, Patrick R.; Plattner, Gian-Kasper; Yohe, Gary W.; Zwiers, Francis W. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. IPCC, 2010.
- [S08] National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press, 2019. DOI: 10.17226/25303.
- [S09] Draper, David. “Assessment and Propagation of Model Uncertainty.” *Journal of the Royal Statistical Society: Series B* 57(1), 1995, 45–70. DOI: 10.1111/j.2517-6161.1995.tb02015.x.
- [S10] Possolo, Antonio; Gelabert, Maria Victoria; Hibbert, David; Stohner, Jürgen; Bodnar, Olha; Meija, Juris. *Guía Breve sobre la Incertidumbre de Medición*. NIST Technical Note 2330, 2025. DOI: 10.6028/NIST.TN.2330.
- [S11] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement. Amendment 1: Nonlinearity in measurement models*. JCGM 100:2008/Amd.1:2026. DOI: 10.59161/PPDI3267.

KNO-016 · Cómo se construye la evidencia: integrando las herramientas

KNO · Formas de conocer

Juan Segurola

v0.6 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia al publicar:** Todos los derechos reservados

Resumen del artículo

La evidencia no es una sustancia que aparezca automáticamente cuando se acumulan datos. ¿Qué debe conectarse para que una observación sea relevante para una afirmación científica? Este artículo sigue esa pregunta desde la formulación de afirmaciones y la medición hasta el diseño de estudios, la modelización, la identificación causal, la incertidumbre, la replicación, la reproducibilidad, la triangulación, la directitud, la transparencia y el escrutinio colectivo. Pregunta qué vínculos son empíricos, cuáles dependen de supuestos, cuáles pueden probarse de forma independiente y cuáles siguen siendo vulnerables a errores compartidos. También pregunta por qué más estudios no significan necesariamente evidencia más fuerte si repiten las mismas debilidades y por qué la evidencia puede volverse más fiable sin convertirse en certeza. El objetivo es hacer explícita la arquitectura del apoyo probatorio para que otro lector vea no solo la conclusión propuesta, sino cómo debería sostenerla cada paso.

En 60 segundos

La evidencia científica es un argumento estructurado con anclajes empíricos. Una afirmación gana credibilidad cuando las mediciones son trazables, el diseño aborda las alternativas relevantes, el modelo es apropiado para la pregunta, la incertidumbre se representa honestamente y análisis o estudios independientes exponen la afirmación a modos de fallo diferentes.

Ninguna etiqueta como «revisado por pares», «estadísticamente significativo», «replicado» o «aleatorizado» garantiza corrección. La fuerza de la evidencia depende del ajuste entre pregunta y método, de la magnitud y dirección de los posibles sesgos, de la consistencia entre estudios relevantes, de la precisión, de la directitud y de la transparencia. La idea práctica es separar el resultado observado de los supuestos que lo conectan con la interpretación. Cada herramienta reduce algunas ambigüedades y deja otras abiertas; por eso el análisis debe declarar qué puede identificar, qué depende de condiciones adicionales y qué observación podría contradecir la conclusión. La precisión formal no elimina por sí sola esos límites.

Por qué importa esta pregunta

Los científicos rara vez observan una conclusión directamente. Observan huellas, medidas o resultados y los conectan con una afirmación mediante diseño, modelos, supuestos y conocimiento de fondo. Cuando esos vínculos quedan implícitos, el desacuerdo es difícil de diagnosticar y los fallos de replicación son difíciles de interpretar. Hacer explícita la cadena probatoria aclara dónde un resultado es fuerte, dónde depende de supuestos y qué nueva observación podría cambiarlo. Eso importa para la práctica científica y para decidir cuánta confianza merece una afirmación.

1. La evidencia empieza con una afirmación suficientemente precisa para poder fallar

Antes de preguntar si la evidencia es fuerte, hay que saber qué afirmación se está apoyando.

[C01] Una afirmación científica debe identificar su objetivo con suficiente claridad para que observaciones o análisis puedan contar en su contra. Las narrativas amplias que cambian de significado después de cada resultado son difíciles de someter a prueba porque el objetivo evidencial se desplaza. [S01, S02]

Las definiciones operacionales no son mero trámite. Conectan el lenguaje teórico con procedimientos observables. Una misma palabra puede referirse a estimandos, mediciones o clasificaciones diferentes. Si el objetivo cambia entre diseño, análisis e interpretación, la evidencia aparente puede responder a una pregunta distinta de la planteada inicialmente.

En qué punto estamos

Antes de preguntar si la evidencia es fuerte, hay que saber qué afirmación se está apoyando. Las definiciones operacionales no son mero trámite. Conectan el lenguaje teórico con procedimientos observables. Una misma palabra puede referirse a estimandos, mediciones o clasificaciones diferentes. La sección siguiente examina la idea de que las observaciones se producen, no simplemente se encuentran.

2. Las observaciones se producen, no simplemente se encuentran

KNO-008 mostró que los instrumentos amplían los sentidos; KNO-011 mostró que los números dependen de convenciones de medida. Aquí esas lecciones pasan a formar parte de la cadena de evidencia.

[C02] El valor evidencial de una medición depende de cómo se definió el mensurando o variable, cómo se calibró el instrumento o procedimiento, qué correcciones se aplicaron y qué incertidumbre permanece. Una salida numérica no se valida a sí misma. [S03]

Una fotografía, un espectro, un ensayo analítico o una respuesta de encuesta ya son el resultado de decisiones sobre muestreo, resolución, procesamiento y clasificación. La evidencia empieza antes del análisis estadístico.

En qué punto estamos

KNO-008 mostró que los instrumentos amplían los sentidos; KNO-011 mostró que los números dependen de convenciones de medida. Aquí esas lecciones pasan a formar parte de la cadena de evidencia. Una fotografía, un espectro, un ensayo analítico o una respuesta de encuesta ya son el resultado de decisiones sobre muestreo, resolución, procesamiento y clasificación. La sección siguiente examina la idea de que el diseño determina qué alternativas pueden excluirse.

3. El diseño determina qué alternativas pueden excluirse

Preguntas distintas requieren diseños distintos.

[C03] Controles experimentales, aleatorización, cegamiento, medición repetida y estructuras comparativas observacionales abordan sesgos potenciales diferentes. Ninguno es un marcador universal de calidad independiente de la pregunta científica. [S04]

Para preguntas causales, el diseño determina qué contrastes contrafactuales pueden identificarse plausiblemente. Para preguntas descriptivas, un muestreo representativo puede importar más que la asignación aleatoria de tratamientos. En ciencias históricas, huellas convergentes y consistencia mecánica pueden soportar la evidencia cuando intervenir es imposible.

En qué punto estamos

Preguntas distintas requieren diseños distintos. Para preguntas descriptivas, un muestreo representativo puede importar más que la asignación aleatoria de tratamientos. En ciencias históricas, huellas convergentes y consistencia mecánica pueden soportar la evidencia cuando intervenir es imposible. El diseño limita qué explicaciones siguen siendo compatibles. La sección siguiente examina la idea de que los modelos son puentes entre observaciones y afirmaciones.

4. Los modelos son puentes entre observaciones y afirmaciones

Los datos rara vez hablan directamente de la proposición de interés. Un modelo traduce observaciones en cantidades, mecanismos o predicciones.

[C04] Toda afirmación evidencial basada en modelos está condicionada en algún grado por elecciones de modelización. Una buena práctica distingue lo directamente observado de lo inferido a través del modelo y prueba la sensibilidad a alternativas plausibles. [S06]

No es un defecto exclusivo de las simulaciones complejas. Incluso un ajuste lineal, una curva de calibración o un umbral diagnóstico constituyen modelos.

En qué punto estamos

Los datos rara vez hablan directamente de la proposición de interés. No es un defecto exclusivo de las simulaciones complejas. Incluso un ajuste lineal, una curva de calibración o un umbral diagnóstico constituyen modelos. Lo importante aquí es la relación declarada entre evidencia, método e inferencia, no una promesa genérica de certeza. La sección siguiente examina la idea de que la evidencia causal exige más que asociación.

5. La evidencia causal exige más que asociación

KNO-014 fijó la frontera causal central.

[C05] Una asociación observada puede apoyar una afirmación causal solo mediante supuestos y características del diseño que aborden confusión, causalidad inversa, selección y otras explicaciones competidoras. Por tanto, la fuerza de la evidencia causal depende de la estrategia de identificación, no solo del tamaño o significación estadística de la asociación. [S04, S07]

Pueden combinarse métodos causales distintos, pero múltiples análisis que compartan el mismo sesgo oculto no constituyen confirmación independiente.

En qué punto estamos

KNO-014 fijó la frontera causal central. Pueden combinarse métodos causales distintos, pero múltiples análisis que compartan el mismo sesgo oculto no constituyen confirmación independiente. La conclusión resulta útil solo en el papel definido por esta sección, manteniendo visibles sus condiciones probatorias. La sección siguiente examina la idea de que la incertidumbre pertenece dentro de la afirmación.

6. La incertidumbre pertenece dentro de la afirmación

KNO-015 trató la incertidumbre como objeto metodológico y no como un intervalo decorativo.

[C06] La evaluación de evidencia debe distinguir incertidumbre de medida, muestreo, modelización y otras fuentes reconocidas. Un intervalo estrecho dentro de un modelo no establece que el error sistemático no modelizado sea pequeño. [S03, S08]

Una afirmación fuerte especifica por ello no solo una estimación, sino también las condiciones bajo las cuales se espera que sea fiable.

En qué punto estamos

KNO-015 trató la incertidumbre como objeto metodológico y no como un intervalo decorativo. Una afirmación fuerte especifica por ello no solo una estimación, sino también las condiciones bajo las cuales se espera que sea fiable. Su valor está en hacer inspeccionable este paso inferencial concreto, no en tratar el método como autojustificado. La sección siguiente examina la idea de que replicación y reproducibilidad prueban vínculos diferentes.

7. Replicación y reproducibilidad prueban vínculos diferentes

Los términos se confunden con frecuencia.

[C07] Las National Academies distinguen reproducibilidad computacional, obtener resultados computacionales consistentes usando los mismos datos, código y métodos, de replicabilidad, obtener resultados consistentes en estudios que abordan la misma pregunta científica con nuevos datos. [S01]

[C08] Una replicación exitosa aumenta la confianza pero no demuestra que una afirmación original sea universalmente correcta; una replicación fallida tampoco la refuta por sí sola de forma concluyente, porque las diferencias pueden proceder de heterogeneidad real, variables no controladas, defectos metodológicos u otras causas. [S01]

Esta asimetría importa. La repetición aporta evidencia sobre robustez, no funciona como un operador lógico de demostración.

En qué punto estamos

Los términos se confunden con frecuencia. La repetición aporta evidencia sobre robustez, no funciona como un operador lógico de demostración. Lo importante aquí es la relación declarada entre evidencia, método e inferencia, no una promesa genérica de certeza. La sección siguiente examina la idea de que la triangulación es más fuerte cuando los errores son distintos.

8. La triangulación es más fuerte cuando los errores son distintos

Evidencia independiente no significa simplemente «más estudios».

[C09] La triangulación gana fuerza cuando métodos diferentes abordan la misma pregunta y poseen fuentes clave de sesgo diferentes. La convergencia es más informativa cuando no se espera que los sesgos de todos los enfoques empujen el resultado en la misma dirección. [S07]

Diez estudios con el mismo instrumento, mecanismo de selección y modelo pueden reproducir un sesgo compartido. Un conjunto menor con diseños genuinamente diferentes puede tener mayor valor diagnóstico.

En qué punto estamos

Evidencia independiente no significa simplemente «más estudios». Un conjunto menor con diseños genuinamente diferentes puede tener mayor valor diagnóstico. La conclusión resulta útil solo en el papel definido por esta sección, manteniendo visibles sus condiciones probatorias. La sección siguiente examina la idea de que la evidencia es acumulativa, pero acumular exige estructura.

9. La evidencia es acumulativa, pero acumular exige estructura

Meta-análisis, revisiones sistemáticas y marcos de evidencia pueden combinar estudios, aunque la combinación solo tiene sentido cuando las preguntas subyacentes son suficientemente compatibles.

[C10] Marcos como GRADE examinan explícitamente riesgo de sesgo, inconsistencia, indirectitud, imprecisión y sesgo de difusión/publicación en lugar de tratar por sí solo el número de estudios como certeza. [S09]

No se afirma que GRADE gobierne toda la ciencia. Ilustra una arquitectura general: los cuerpos de evidencia deben evaluarse por error sistemático, coherencia, relevancia y precisión, no solo agregarse.

En qué punto estamos

Meta-análisis, revisiones sistemáticas y marcos de evidencia pueden combinar estudios, aunque la combinación solo tiene sentido cuando las preguntas subyacentes son suficientemente compatibles. No se afirma que GRADE gobierne toda la ciencia. Ilustra una arquitectura general: los cuerpos de evidencia deben evaluarse por error sistemático, coherencia, relevancia y precisión, no solo agregarse. La sección siguiente examina la idea de que la directitud importa.

10. La directitud importa

Una evidencia puede ser exacta y, aun así, responder a la pregunta equivocada.

Un sustituto de laboratorio puede no predecir un resultado clínico. Un resultado en una población puede no transportarse a otra. Una variable proxy puede medir solo una parte del constructo previsto.

[C11] El concepto de indirectitud formaliza desajustes entre la pregunta de interés y las poblaciones, intervenciones o exposiciones, comparadores y resultados realmente estudiados. [S09, S10]

La validez externa y la transportabilidad pertenecen, por tanto, a la evaluación de evidencia y no a un añadido posterior a la validez interna.

En qué punto estamos

Una evidencia puede ser exacta y, aun así, responder a la pregunta equivocada. Un sustituto de laboratorio puede no predecir un resultado clínico. Un resultado en una población puede no transportarse a otra. Una variable proxy puede medir solo una parte del constructo previsto. La sección siguiente examina la idea de que la transparencia cambia lo que otros pueden comprobar.

11. La transparencia cambia lo que otros pueden comprobar

La evidencia es más fuerte cuando puede inspeccionarse el camino desde el material bruto hasta la conclusión.

[C12] Una comunicación clara de métodos, pasos de procesamiento de datos, elecciones analíticas e incertidumbre facilita la comprobación computacional y ayuda a otros investigadores a identificar errores o probar robustez. La reproducibilidad se favorece mediante acceso a los artefactos digitales pertinentes y métodos suficientemente completos. [S01, S02]

La transparencia no vuelve correcto un análisis defectuoso. Hace que sus defectos sean más descubribles.

En qué punto estamos

La evidencia es más fuerte cuando puede inspeccionarse el camino desde el material bruto hasta la conclusión. La transparencia no vuelve correcto un análisis defectuoso. Hace que sus defectos sean más descubribles. La inspección externa depende de esa trazabilidad. La sección siguiente examina la idea de que revisión por pares y consenso son filtros, no sustitutos de la evidencia.

12. Revisión por pares y consenso son filtros, no sustitutos de la evidencia

La revisión por pares puede detectar errores, pedir aclaraciones y mejorar la comunicación, pero no reproduce de forma independiente el experimento subyacente. El consenso puede resumir un panorama evidencial maduro, pero unos supuestos compartidos también pueden propagarse por un campo.

Estos procesos institucionales pertenecen plenamente a KNO-020. Aquí la frontera es más simple: la validación social puede reforzar el control de calidad, pero la cadena evidencial debe seguir conectando observaciones y afirmaciones.

En qué punto estamos

La revisión por pares puede detectar errores, pedir aclaraciones y mejorar la comunicación, pero no reproduce de forma independiente el experimento subyacente. El consenso puede resumir un panorama evidencial maduro, pero unos supuestos compartidos también pueden propagarse por un campo. Estos procesos institucionales pertenecen plenamente a KNO-020. La sección siguiente examina una auditoría práctica de evidencia.

13. Una auditoría práctica de evidencia

Antes de aceptar una conclusión científica fuerte, preguntar:

1. **¿Cuál es exactamente la afirmación?** ¿Descriptiva, predictiva, causal, mecanística o histórica?
2. **¿Qué observaciones la anclan?** ¿Qué se midió o registró directamente?
3. **¿Cómo se construyeron las variables?** ¿Qué clasificaciones, umbrales o transformaciones intervinieron?
4. **¿Qué diseño aborda las alternativas?** ¿Controles, aleatorización, emparejamiento, orden temporal, variación natural, casos comparativos?
5. **¿Qué modelo conecta datos y afirmación?** ¿Qué supuestos son esenciales?
6. **¿Qué incertidumbres están cuantificadas?** ¿Cuáles siguen reconocidas pero no cuantificadas?
7. **¿Qué sesgos podrían desplazar la respuesta?** ¿En qué dirección?
8. **¿Se ha reproducido el cálculo?** ¿Puede ejecutarse de nuevo el camino analítico?
9. **¿Se ha replicado o triangulado el hallazgo?** ¿Los nuevos modos de fallo son realmente distintos?
10. **¿Qué tan directa es la evidencia para esta población, cantidad o mecanismo objetivo?**
11. **¿Es consistente con un cuerpo de evidencia más amplio?** Si no, ¿es informativo el desacuerdo?
12. **¿Qué conclusión está realmente justificada?** No convertir «compatible con» en «demostrado».

En qué punto estamos

Antes de aceptar una conclusión científica fuerte, preguntar: ¿Qué incertidumbres están cuantificadas? ¿Cuáles siguen reconocidas pero no cuantificadas? ¿Qué observaciones la anclan? ¿Qué se midió o registró directamente? ¿Cómo se construyeron las variables? ¿Qué clasificaciones, umbrales o transformaciones intervinieron? La sección siguiente examina la idea de que la evidencia puede fortalecerse sin convertirse en certeza.

14. La evidencia puede fortalecerse sin convertirse en certeza

La ciencia acumula restricciones. Las mediciones estrechan valores posibles; los experimentos eliminan alternativas; los modelos organizan implicaciones; métodos independientes prueban vulnerabilidades distintas; la replicación explora robustez; la síntesis muestra dónde los resultados convergen y dónde no.

El punto final rara vez es la certeza. Es una afirmación cuyas alternativas restantes resultan progresivamente más difíciles de sostener.

En qué punto estamos

Las mediciones estrechan valores posibles; los experimentos eliminan alternativas; los modelos organizan implicaciones; métodos independientes prueban vulnerabilidades distintas; la replicación explora robustez; la síntesis muestra dónde los resultados convergen y dónde no. Esto delimita el alcance del artículo y prepara la síntesis final sin ampliar la afirmación más allá de lo establecido en las secciones anteriores.

Nota de frontera

KNO-016 es deliberadamente un **artículo de síntesis**. No sustituye los métodos detallados de KNO-008 a KNO-015 y no se adelanta a KNO-020 sobre revisión y consenso ni a KNO-021 sobre ciencia abierta. Su función es mostrar cómo medición, diseño experimental, modelización, causalidad, incertidumbre, reproducibilidad, replicación y triangulación se combinan en un argumento evidencial auditable.

Lo que sabemos y cómo lo sabemos

La evidencia se construye conectando afirmaciones precisas con observaciones mediante medida, diseño, modelos y supuestos inferenciales. Replicación, reproducibilidad, triangulación, directitud y transparencia refuerzan vínculos distintos, sobre todo cuando sus modos de fallo no son compartidos.

Lo que todavía no sabemos

Ninguna auditoría garantiza que se hayan imaginado todas las alternativas relevantes, que todas las fuentes de sesgo sean independientes o que un modelo bien apoyado sea la única explicación posible. La evidencia puede fortalecerse sin convertirse en conclusión lógica final.

Qué podría hacernos cambiar de opinión

La evaluación probatoria debe cambiar si falla la calibración de una medida clave, el diseño permite una explicación alternativa, se rompe una predicción del modelo, una replicación comparable fracasa o un método con estructura de error diferente contradice el resultado existente.

Si solo recuerdas una cosa

La evidencia fuerte no es una pila de resultados, sino una cadena de apoyo cuyos supuestos, errores y pruebas pueden inspeccionarse.

Artículos relacionados

- KNO-012 · Experiments and Variable Control
- KNO-013 · Models, Maps and Reality
- KNO-014 · Causality and Correlation

- KNO-015 · Uncertainty and Probability
- KNO-020 · Review, Consensus and Disagreement

Referencias

- [S01] National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press, 2019. DOI: 10.17226/25303.
- [S02] National Academies of Sciences, Engineering, and Medicine. “Understanding Reproducibility and Replicability,” chapter 3 in *Reproducibility and Replicability in Science*, pp. 39–54. Washington, DC: The National Academies Press, 2019. DOI: 10.17226/25303.
- [S03] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. JCGM 100:2008(E). DOI: 10.59161/JCGM100-2008E.
- [S04] Hernán, Miguel A. “A Definition of Causal Effect for Epidemiological Research.” *Journal of Epidemiology and Community Health* 58(4), 2004, 265–271. DOI: 10.1136/jech.2002.006361.
- [S05] Schulz, Kenneth F.; Altman, Douglas G.; Moher, David; CONSORT Group. “CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.” *BMJ* 340, 2010, c332. DOI: 10.1136/bmj.c332.
- [S06] Frigg, Roman; Hartmann, Stephan. “Models in Science.” *The Stanford Encyclopedia of Philosophy*, edición de verano de 2025; revisión sustantiva del 2 de abril de 2025. <https://plato.stanford.edu/archives/sum2025/entries/models-science/>
- [S07] Lawlor, Debbie A.; Tilling, Kate; Davey Smith, George. “Triangulation in aetiological epidemiology.” *International Journal of Epidemiology* 45(6), 2016, 1866–1886. DOI: 10.1093/ije/dyw314.
- [S08] Mastrandrea, Michael D.; Field, Christopher B.; Stocker, Thomas F.; Edenhofer, Ottmar; Ebi, Kristie L.; Frame, David J.; Held, Hermann; Kriegler, Elmar; Mach, Katharine J.; Matschoss, Patrick R.; Plattner, Gian-Kasper; Yohe, Gary W.; Zwiers, Francis W. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. IPCC, 2010. <https://www.ipcc.ch/site/assets/uploads/2018/05/uncertainty-guidance-note.pdf>
- [S09] Neumann, Ignacio; Brennan, Sue; Meerpohl, Joerg; Davoli, Marina; Alonso Coello, Pablo; Akl, Elie; Dahm, Philipp; Skoetz, Nicole; Xia, Jun; Brozek, Jan; Schünemann, Holger J. “Overview of the GRADE approach.” *GRADE Book*. Last modified 12 May 2026; consulted 5 September 2026. <https://book.gradeapro.org/guideline/overview-of-the-grade-approach>
- [S10] “Indirectness.” *GRADE Book*. Last modified 12 May 2026; consulted 5 September 2026. <https://book.gradeapro.org/guideline/indirectness>. Author display currently contains an apparent duplicated Holger J. Schünemann entry; author chain therefore remains P pending human confirmation and is not normalised in this bibliography.

KNO-017 · ¿Son las matemáticas un lenguaje?

KNO · Formas de conocer

Juan Segurola

v0.6 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia al publicar:** Todos los derechos reservados

Resumen del artículo

A menudo se dice que las matemáticas son un lenguaje, pero ¿qué se está comparando exactamente cuando se afirma eso? Este artículo pregunta si la notación matemática, los sistemas formales, las prácticas de demostración, la abstracción y la modelización aplicada cumplen las mismas funciones que el lenguaje natural, o si la analogía solo funciona para algunas partes de las matemáticas. Examina sintaxis, semántica, interpretación, demostración, cálculo y representación cognitiva externa separando las expresiones matemáticas de las prácticas más amplias en las que se crean y justifican. También pregunta por qué los símbolos formales permiten manipulaciones extraordinariamente precisas sin portar significado por sí solos y por qué la eficacia de las matemáticas en ciencia no resuelve qué son en último término. La cuestión no es, por tanto, si las matemáticas usan lenguajes, sino hasta dónde puede llevarse la analogía antes de ocultar diferencias importantes.

En 60 segundos

Una expresión matemática tiene sintaxis: los símbolos aparecen en disposiciones permitidas. En un sistema formal, unas reglas pueden determinar qué expresiones están bien formadas y qué inferencias están autorizadas. Eso hace que algunas partes de las matemáticas se parezcan a un lenguaje en un sentido técnico preciso. Pero los símbolos no llevan incorporada por sí solos su interpretación científica. Una misma estructura formal puede interpretarse en dominios diferentes, y una ecuación científica solo adquiere contenido empírico cuando cantidades, unidades, mediciones y reglas de correspondencia la conectan con observaciones. Las demostraciones matemáticas ordinarias tampoco suelen escribirse como derivaciones completamente formales. Mezclan símbolos con prosa, diagramas, convenciones y pasos omitidos que los lectores formados pueden reconstruir.

Llamar a las matemáticas «un lenguaje» es, por tanto, útil si significa que proporcionan sistemas de representación con sintaxis, reglas inferenciales y convenciones aprendidas. Se vuelve engañoso si implica que las matemáticas son simplemente otra lengua natural, que la notación constituye significado por sí misma o que la eficacia de las matemáticas demuestra que la realidad literalmente habla matemáticas.

Por qué importa esta pregunta

Llamar lenguaje a las matemáticas puede iluminar cómo notación, sintaxis y convenciones compartidas hacen portable el razonamiento, pero la metáfora también puede ocultar diferencias entre manipulación formal, demostración, significado e interpretación empírica. La cuestión importa porque las matemáticas son centrales en casi toda ciencia cuantitativa y sus símbolos parecen a menudo más autosuficientes de lo que son. Entender qué aporta la notación y qué no puede aportar por sí sola ayuda a separar validez matemática de adecuación del modelo, cálculo de explicación y precisión formal de afirmaciones sobre el mundo.

1. Hay más de un sentido de «lenguaje»

Las lenguas naturales combinan sintaxis, semántica, pragmática y convención social. Los lenguajes formales aíslan deliberadamente algunas de estas características. En lógica clásica, un lenguaje formal puede especificarse como un conjunto recursivamente definido de cadenas sobre un alfabeto, mientras que las reglas deductivas y la semántica aportan estructura inferencial o interpretativa [S01].

Las matemáticas utilizan muchos sistemas formales o semiformales de este tipo, pero también recurren al lenguaje natural y a los diagramas. Por eso, la frase «las matemáticas son un lenguaje» no constituye una única tesis técnica. Puede referirse a la notación, a la sintaxis formal, a la demostración, a la comunicación o a la representación.

[C01] Partes de las matemáticas pueden representarse como lenguajes formales con sintaxis y reglas inferenciales explícitas, pero eso no establece por sí solo que toda la práctica matemática sea un único lenguaje formal. [S01,S02]

En qué punto estamos

Las lenguas naturales combinan sintaxis, semántica, pragmática y convención social. En lógica clásica, un lenguaje formal puede especificarse como un conjunto recursivamente definido de cadenas sobre un alfabeto, mientras que las reglas deductivas y la semántica aportan estructura inferencial o interpretativa [S01]. Los lenguajes formales aíslan deliberadamente algunas de estas características. La sección siguiente examina la idea de que la sintaxis todavía no es significado.

2. La sintaxis todavía no es significado

Una cadena puede estar bien formada sin especificar todavía qué denotan sus símbolos. Los textos de lógica distinguen sintaxis y semántica precisamente porque una expresión formal y una interpretación de esa expresión son objetos diferentes [S01].

Esto importa científicamente. La ecuación $y = ax + b$ puede describir muchos sistemas, o ninguno, dependiendo de cómo se hagan corresponder sus símbolos con cantidades y condiciones. La validez formal, por tanto, no certifica la interpretación empírica.

[C02] Una sintaxis matemática bien formada no determina por sí sola el significado empírico de los símbolos manipulados. [S01]

En qué punto estamos

Una cadena puede estar bien formada sin especificar todavía qué denotan sus símbolos. La ecuación $y = ax + b$ puede describir muchos sistemas, o ninguno, dependiendo de cómo se hagan corresponder sus símbolos con cantidades y condiciones. La sección siguiente examina la idea de que la notación cambia lo que puede verse y manipularse.

3. La notación cambia lo que puede verse y manipularse

La notación no es un envoltorio transparente de un pensamiento que ya existía intacto. Una buena notación hace que ciertas estructuras sean más fáciles de comparar, transformar y reutilizar. Los estudios históricos sobre notación matemática muestran cómo la compresión simbólica puede poner de manifiesto patrones que resultan engorrosos de manejar en prosa [S03].

Esto no significa que la notación cree la verdad matemática. Significa que las representaciones modifican la accesibilidad cognitiva y práctica de las relaciones. La notación algebraica, los sistemas de coordenadas y la notación de operadores hacen que unas operaciones sean fáciles y otras costosas de realizar.

[C03] La notación matemática puede reorganizar la accesibilidad de relaciones y operaciones incluso cuando el contenido matemático subyacente no cambia. [S03]

En qué punto estamos

La notación no es un envoltorio transparente de un pensamiento que ya existía intacto. Los estudios históricos sobre notación matemática muestran cómo la compresión simbólica puede poner de manifiesto patrones que resultan engorrosos de manejar en prosa [S03]. La sección siguiente examina la idea de que la formalización hace inspeccionables las reglas.

4. La formalización hace inspeccionables las reglas

La formalización sustituye muchas convenciones tácitas por gramáticas, definiciones y reglas de inferencia explícitas. La lógica matemática hizo posible estudiar las propias teorías matemáticas como objetos formales [S02,S04]. Los asistentes de demostración modernos extienden computacionalmente esa idea: definiciones y pruebas pueden codificarse en un lenguaje formal que una máquina analiza y contrasta con un sistema fundacional [S05].

La ganancia es auditabilidad, no omnisciencia. La verificación formal comprueba una derivación con respecto a las definiciones, axiomas, bibliotecas y entorno de software codificados. No decide si una formalización científica eligió el fenómeno correcto que debía representar.

[C04] La formalización puede hacer mecánicamente inspeccionables los compromisos sintácticos e inferenciales, mientras deja la elección y la interpretación de la formalización como un problema separado. [S04,S05]

En qué punto estamos

La formalización sustituye muchas convenciones tácitas por gramáticas, definiciones y reglas de inferencia explícitas. La ganancia es auditabilidad, no omnisciencia. Los asistentes de demostración modernos extienden computacionalmente esa idea: definiciones y pruebas pueden codificarse en un

lenguaje formal que una máquina analiza y contrasta con un sistema fundacional [S05]. La sección siguiente examina la idea de que una demostración matemática es más que una cadena impresa de símbolos.

5. Una demostración matemática es más que una cadena impresa de símbolos

La práctica matemática suele comunicar las demostraciones en una forma socialmente comprimida. Los argumentos publicados omiten transformaciones elementales, descansan en definiciones compartidas y a menudo mezclan prosa, ecuaciones y diagramas. Por ello, la filosofía contemporánea de la práctica matemática distingue las derivaciones formales de las demostraciones informales o semiformales que los matemáticos escriben y evalúan realmente [S06].

Una demostración puede tener, en consecuencia, dos criterios relacionados: que su conclusión se siga conforme a las reglas matemáticas aceptadas y que la presentación haga inteligible y comprobable para su audiencia el razonamiento relevante.

[C05] Las demostraciones matemáticas ordinarias suelen presentarse en lenguaje natural intercalado con símbolos formales y no como derivaciones formales completamente expandidas. [S06]

En qué punto estamos

La práctica matemática suele comunicar las demostraciones en una forma socialmente comprimida. Una demostración puede tener, en consecuencia, dos criterios relacionados: que su conclusión se siga conforme a las reglas matemáticas aceptadas y que la presentación haga inteligible y comprobable para su audiencia el razonamiento relevante. La sección siguiente examina la idea de que un lenguaje formal puede tener varias interpretaciones.

6. Un lenguaje formal puede tener varias interpretaciones

Una razón por la que los sistemas formales son reutilizables es que una misma sintaxis puede admitir distintos modelos o interpretaciones. Es una fortaleza, pero impide identificar de manera simple una ecuación con una única realidad física.

En ciencia aplicada, una estructura matemática solo se convierte en modelo de un sistema objetivo después de especificar variables, parámetros, condiciones de contorno y relaciones de correspondencia [S01,S07]. KNO-013 trata la adecuación de los modelos; aquí el punto es más estrecho y representacional: la forma matemática subdetermina su aplicación.

[C06] La reutilización de una misma estructura matemática entre dominios depende de la interpretación; la estructura formal por sí sola no selecciona su objetivo científico. [S01,S07]

En qué punto estamos

Una razón por la que los sistemas formales son reutilizables es que una misma sintaxis puede admitir distintos modelos o interpretaciones. Es una fortaleza, pero impide identificar de manera simple una ecuación con una única realidad física. La interpretación sigue siendo necesaria. La sección siguiente examina la idea de que la abstracción es una capacidad, no un defecto.

7. La abstracción es una capacidad, no un defecto

Las matemáticas suelen suprimir detalle material para preservar estructura. Un mismo grupo, grafo, ecuación diferencial o espacio de probabilidad puede ser instanciado por sistemas de composición física muy diferente. Esa abstracción favorece la transferencia de métodos entre dominios.

Pero la abstracción también crea una puerta de control: los rasgos omitidos deben ser verdaderamente irrelevantes para la pregunta planteada. La derivación matemática puede ser exacta dentro de la abstracción mientras la abstracción resulta científicamente inadecuada.

[C07] La abstracción matemática favorece la transferencia al preservar estructura seleccionada entre dominios diferentes, pero la adecuación científica depende de que los rasgos omitidos sean irrelevantes para la pregunta objetivo. [S02,S07]

En qué punto estamos

Las matemáticas suelen suprimir detalle material para preservar estructura. Un mismo grupo, grafo, ecuación diferencial o espacio de probabilidad puede ser instanciado por sistemas de composición física muy diferente. Esa abstracción favorece la transferencia de métodos entre dominios. La sección siguiente examina la idea de que las matemáticas pueden expresar relaciones con mayor precisión que la prosa.

8. Las matemáticas pueden expresar relaciones con mayor precisión que la prosa

El lenguaje natural puede decir que una cantidad aumenta con otra; una ecuación puede especificar la relación funcional, la dependencia de parámetros y los casos límite. También pueden hacerse explícitas las unidades, los dominios y los supuestos. Esa precisión es una de las razones por las que las matemáticas son indispensables en las ciencias cuantitativas maduras.

Sin embargo, la precisión de expresión no debe confundirse con la precisión del conocimiento. Un modelo puede contener muchos dígitos significativos mientras sus parámetros están mal medidos o su estructura es incorrecta. KNO-015 establece la frontera correspondiente a la incertidumbre.

[C08] La representación matemática puede aumentar la precisión sobre relaciones y supuestos sin garantizar que las cantidades representadas o el modelo sean empíricamente exactos. [S07]

En qué punto estamos

El lenguaje natural puede decir que una cantidad aumenta con otra; una ecuación puede especificar la relación funcional, la dependencia de parámetros y los casos límite. Sin embargo, la precisión de expresión no debe confundirse con la precisión del conocimiento. La sección siguiente examina la idea de que calcular y explicar no son lo mismo.

9. Calcular y explicar no son lo mismo

Una derivación puede producir un resultado numérico correcto sin explicar por qué ocurre el fenómeno. A la inversa, un argumento explicativo puede depender de una estructura cualitativa y no de un cálculo largo. Por ello, filósofos de las matemáticas y de la ciencia distinguen demostración, cálculo, predicción y explicación en lugar de tratarlos como sinónimos [S06,S07].

La distinción importa para la comunicación. «La ecuación da el número correcto» es evidencia de adecuación empírica en un dominio comprobado, no automáticamente una explicación del mecanismo.

[C09] Cálculo correcto, demostración válida, predicción acertada y explicación científica son logros epistémicos relacionados pero no idénticos. [S06,S07]

En qué punto estamos

Una derivación puede producir un resultado numérico correcto sin explicar por qué ocurre el fenómeno. A la inversa, un argumento explicativo puede depender de una estructura cualitativa y no de un cálculo largo. Por ello, filósofos de las matemáticas y de la ciencia distinguen demostración, cálculo, predicción y explicación en lugar de tratarlos como sinónimos [S06,S07]. La sección siguiente pregunta por qué las matemáticas funcionan tan bien en ciencia.

10. ¿Por qué funcionan tan bien las matemáticas en ciencia?

Wigner planteó célebremente la eficacia de las matemáticas en las ciencias naturales como algo que, en sí mismo, exige explicación [S08]. Su ensayo es históricamente importante porque identifica un problema real: estructuras matemáticas desarrolladas con una autonomía considerable encuentran a menudo, más tarde, aplicaciones científicas sorprendentes.

Pero esa observación no decide qué son las matemáticas. No demuestra el platonismo, el formalismo, el estructuralismo ni la afirmación de que el universo literalmente «sea matemáticas». Esas son posiciones filosóficas que requieren argumentos adicionales.

[C10] La aplicación exitosa de las matemáticas a la ciencia natural es un hecho empírico y filosófico que debe explicarse, no una prueba por sí sola de una única ontología de las matemáticas. [S02,S08]

En qué punto estamos

Wigner planteó célebremente la eficacia de las matemáticas en las ciencias naturales como algo que, en sí mismo, exige explicación [S08]. Pero esa observación no decide qué son las matemáticas. No demuestra el platonismo, el formalismo, el estructuralismo ni la afirmación de que el universo literalmente «sea matemáticas». La sección siguiente examina las matemáticas como infraestructura cognitiva externa.

11. Las matemáticas como infraestructura cognitiva externa

Las matemáticas escritas permiten que el razonamiento persista fuera de una única memoria de trabajo. Las expresiones pueden copiarse, transformarse, comprobarse y compartirse. Una notación puede conservar estados intermedios de modo que una inferencia larga no tenga que reconstruirse mentalmente desde cero [S03,S05].

Esto conecta KNO-017 con KNO-022, que tratará la computación como memoria externa de forma más general. Aquí la afirmación más estrecha es que las matemáticas simbólicas constituyen una tecnología externa de representación además de una disciplina abstracta.

[C11] Las inscripciones matemáticas y las representaciones formales pueden funcionar como estados externos y revisables dentro de cadenas extendidas de razonamiento. [S03,S05]

En qué punto estamos

Las matemáticas escritas permiten que el razonamiento persista fuera de una única memoria de trabajo. Una notación puede conservar estados intermedios de modo que una inferencia larga no tenga que reconstruirse mentalmente desde cero [S03,S05]. Aquí la afirmación más estrecha es que las matemáticas simbólicas constituyen una tecnología externa de representación además de una disciplina abstracta. La sección siguiente sintetiza la respuesta a la pregunta central del artículo.

12. Respuesta: las matemáticas contienen lenguajes, pero no se agotan en ellos

La respuesta más fuerte que puede defenderse es, por tanto, condicional. Las matemáticas contienen lenguajes formales y dependen en gran medida de convenciones simbólicas aprendidas. Estos lenguajes proporcionan representación compacta y estructura inferencial explícita. Pero las matemáticas tal como se practican también incluyen interpretación, razonamiento informal, diagramas, ejemplos, estrategias de demostración y juicios sobre relevancia.

Llamar a las matemáticas «un lenguaje» es útil cuando resalta la representación y la inferencia gobernada por reglas. Es demasiado fuerte cuando reduce todas las matemáticas a sintaxis o trata los símbolos como si se interpretaran a sí mismos.

[C12] «Las matemáticas son un lenguaje» es defendible como metáfora metodológica de sus sistemas representacionales y gobernados por reglas, pero no como descripción exhaustiva de la práctica matemática. [S01,S02,S06]

En qué punto estamos

La respuesta más fuerte que puede defenderse es, por tanto, condicional. Pero las matemáticas tal como se practican también incluyen interpretación, razonamiento informal, diagramas, ejemplos, estrategias de demostración y juicios sobre relevancia. Estos lenguajes proporcionan representación compacta y estructura inferencial explícita. Es demasiado fuerte cuando reduce todas las matemáticas a sintaxis o trata los símbolos como si se interpretaran a sí mismos. La sección siguiente delimita el alcance de KNO-017.

13. Frontera de KNO-017

KNO-017 **no** decide si los objetos matemáticos existen independientemente de las mentes, si la verdad matemática se descubre o se inventa, ni qué filosofía de las matemáticas es correcta. Esas preguntas pertenecen a la serie conceptual IDE.

Tampoco trata en profundidad los diagramas y las codificaciones visuales; eso corresponde a KNO-018. Ni aborda la computación como tecnología general de memoria externa; eso corresponde a KNO-022.

El resultado metodológico es más acotado: la notación matemática, los lenguajes formales y la demostración son herramientas que hacen las relaciones representables, transformables, comunicables e inspeccionables, pero su aplicación al mundo requiere interpretación y evidencia más allá de la sintaxis.

En qué punto estamos

KNO-017 no decide si los objetos matemáticos existen independientemente de las mentes, si la verdad matemática se descubre o se inventa, ni qué filosofía de las matemáticas es correcta. Esas preguntas pertenecen a la serie conceptual IDE. Esto delimita el alcance del artículo y prepara la síntesis final sin ampliar la afirmación más allá de lo establecido en las secciones anteriores.

Lo que sabemos y cómo lo sabemos

Las matemáticas utilizan lenguajes formales con sintaxis e interpretación, y la notación puede hacer el razonamiento más inspeccionable y manejable. Sin embargo, demostración, significado, modelización y explicación dependen de prácticas e interpretaciones que no se agotan en cadenas de símbolos.

Lo que todavía no sabemos

La analogía lingüística no resuelve qué son las matemáticas, por qué sus estructuras se aplican con tanta eficacia en ciencia ni si los objetos matemáticos se descubren, se inventan o se describen de otra manera.

Qué podría hacernos cambiar de opinión

El alcance de la analogía debe reconsiderarse si una propuesta explica la notación formal pero no la demostración, la interpretación, la abstracción o la modelización aplicada, o si otro marco captura mejor esas funciones sin tratar toda la matemática como lingüística.

Si solo recuerdas una cosa

Las matemáticas contienen lenguajes y los usan poderosamente, pero el conocimiento matemático es más amplio que los sistemas formales en que se escribe.

Artículos relacionados

- KNO-003 · From Marks to Symbols
- KNO-013 · Models, Maps and Reality
- KNO-016 · How Evidence Is Built: Putting the Tools Together
- KNO-018 · Images, Diagrams and Visualisation

Referencias

[S01] Shapiro, Stewart; Kouri Kissel, Teresa. "Classical Logic." *The Stanford Encyclopedia of Philosophy*, Fall 2022 archive, substantive revision 29 June 2022. <https://plato.stanford.edu/archives/fall2022/entries/logic-classical/>

[S02] Horsten, Leon. "Philosophy of Mathematics." *The Stanford Encyclopedia of Philosophy*, Spring 2022 Edition, substantive revision 25 January 2022. <https://plato.stanford.edu/archives/spr2022/entries/philosophy-mathematics/>

[S03] Mazur, Joseph. *Enlightening Symbols: A Short History of Mathematical Notation and Its Hidden Powers*. Princeton University Press, 2014/2016 online edition. DOI: 10.23943/princeton/9780691173375.001.0001.

[S04] Weir, Alan. "Formalism in the Philosophy of Mathematics." *The Stanford Encyclopedia of Philosophy*, Fall 2024 Edition, substantive revision 20 February 2024. <https://plato.stanford.edu/archives/fall2024/entries/formalism-mathematics/>

[S05] Avigad, Jeremy; Commelin, Johan; Macbeth, Heather; Topaz, Adam. "Anatomy of a Formal Proof." arXiv:2411.11885, 2024. <https://arxiv.org/abs/2411.11885>

[S06] Tanswell, Fenner Stanley. *Mathematical Rigour and Informal Proof*. Cambridge University Press, 2024. Cambridge Elements in the Philosophy of Mathematics.

[S07] Frigg, Roman; Hartmann, Stephan. "Models in Science." *The Stanford Encyclopedia of Philosophy*, Summer 2025 Edition, substantive revision 2 April 2025. <https://plato.stanford.edu/archives/sum2025/entries/models-science/>

[S08] Wigner, Eugene P. "The Unreasonable Effectiveness of Mathematics in the Natural Sciences." *Communications on Pure and Applied Mathematics* 13(1), 1960, 1–14. DOI: 10.1002/cpa.3160130102.

KNO-018 · Imágenes, diagramas y visualización

KNO · Formas de conocer

Juan Segurola

v0.6 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia al publicar:** Todos los derechos reservados

Resumen del artículo

Las imágenes, diagramas y gráficos científicos pueden hacer visibles patrones difíciles de expresar en prosa, pero ¿qué ocurre exactamente cuando los datos se convierten en una representación visual? Este artículo examina diagramas, codificaciones gráficas, imágenes científicas y cadenas de visualización como herramientas relacionadas pero distintas. Pregunta cómo la disposición espacial cambia el trabajo cognitivo, cómo los canales perceptivos afectan al juicio cuantitativo, cómo escala, encuadre y color influyen en la interpretación y cómo la adquisición y el procesamiento median entre un sistema físico y los píxeles que ve el lector. También pregunta cuándo las vistas alternativas revelan dependencia de decisiones arbitrarias, cuándo el procesamiento aclara en lugar de distorsionar y qué procedencia hace falta para que una figura sea auditable. El foco no es si la visualización es útil, sino qué transformaciones debe poder reconstruir un lector antes de tratar un patrón visual como evidencia.

En 60 segundos

Una representación visual puede reducir trabajo cognitivo porque las relaciones espaciales pueden hacer explícitos patrones que la prosa deja implícitos [1]. Pero los canales visuales no tienen la misma precisión: la posición sobre una escala común suele decodificarse con mayor exactitud que el área o el ángulo [2]. Los ejemplos clásicos de Anscombe muestran por qué los resúmenes numéricos pueden ocultar estructuras radicalmente distintas [3]. Sin embargo, un gráfico no es la realidad en bruto. Las decisiones de diseño seleccionan variables, escalas, codificaciones y vistas [4,5]. Los mapas de color pueden introducir fronteras perceptivas falsas o énfasis desiguales [6,7]. Las imágenes científicas también son datos numéricos: el ajuste de contraste, el filtrado, la segmentación y la reconstrucción pueden revelar estructuras útiles, pero también borrar o inventar rasgos aparentes si el flujo de trabajo no se controla y documenta [8,9]. La pregunta correcta no es si una figura resulta atractiva, sino qué transformación la produjo, qué tarea facilita, qué incertidumbre o información suprime y si otro lector puede reconstruir ese recorrido.

Por qué importa esta pregunta

La ciencia moderna depende de intermediarios visuales: imágenes microscópicas, mapas, espectros, simulaciones, diagramas de redes, mapas de calor y gráficos estadísticos. Una figura puede ser el primer lugar donde aparece un valor atípico, un efecto de lote o una relación no lineal. Eso convierte la visualización en parte de la investigación, no solo de la presentación. El mismo poder crea riesgo: una

representación puede exagerar diferencias débiles, ocultar extremos o convertir decisiones de procesamiento en rasgos aparentemente naturales. Tratar la visualización como instrumento epistémico somete sus transformaciones al mismo escrutinio que medidas y modelos.

1. Un diagrama puede cambiar el trabajo necesario

Larkin y Simon mostraron que representaciones informacionalmente equivalentes pueden diferir en eficiencia computacional: un diagrama espacial puede hacer explícitas relaciones que una descripción enunciativa deja por derivar [1]. Esto no significa que los diagramas sean universalmente superiores. Su ventaja depende de la tarea y de qué relaciones hace fáciles de inspeccionar la representación.

En qué punto estamos

Larkin y Simon mostraron que representaciones informacionalmente equivalentes pueden diferir en eficiencia computacional: un diagrama espacial puede hacer explícitas relaciones que una descripción enunciativa deja por derivar [1]. Esto no significa que los diagramas sean universalmente superiores. Su ventaja depende de la tarea y de qué relaciones hace fáciles de inspeccionar la representación. La sección siguiente examina la idea de que ver también es decodificar.

2. Ver también es decodificar

Los gráficos codifican datos en canales visuales. Cleveland y McGill estudiaron experimentalmente con qué precisión los observadores decodifican codificaciones gráficas elementales, mostrando que la elección del canal afecta al juicio cuantitativo [2]. Por tanto, una visualización tiene una estructura de error perceptivo además de una estructura de error estadístico.

En qué punto estamos

Los gráficos codifican datos en canales visuales. Cleveland y McGill estudiaron experimentalmente con qué precisión los observadores decodifican codificaciones gráficas elementales, mostrando que la elección del canal afecta al juicio cuantitativo [2]. Por tanto, una visualización tiene una estructura de error perceptivo además de una estructura de error estadístico. La sección siguiente examina la idea de que los números resumen pueden ocultar estructura.

3. Los números resumen pueden ocultar estructura

Anscombe construyó conjuntos de datos con estadísticas resumen convencionales casi idénticas y diagramas de dispersión sorprendentemente distintos [3]. La lección no es que los gráficos sustituyan a la estadística. Es que diferentes representaciones exponen propiedades diferentes, por lo que los resúmenes numéricos y visuales deben tratarse como comprobaciones complementarias.

En qué punto estamos

Anscombe construyó conjuntos de datos con estadísticas resumen convencionales casi idénticas y diagramas de dispersión sorprendentemente distintos [3]. La lección no es que los gráficos sustituyan a la estadística. Es que diferentes representaciones exponen propiedades diferentes, por lo que los resúmenes numéricos y visuales deben tratarse como comprobaciones complementarias. La sección siguiente examina la idea de que toda vista tiene una tarea.

4. Toda vista tiene una tarea

El marco de Munzner separa abstracción de datos, abstracción de tareas, codificación visual y validación [4]. Ware relaciona asimismo el diseño de las representaciones con propiedades de la percepción humana [5]. Una figura no es simplemente «buena» en abstracto: es adecuada o inadecuada para una pregunta, audiencia, tipo de datos y decisión determinados.

En qué punto estamos

El marco de Munzner separa abstracción de datos, abstracción de tareas, codificación visual y validación [4]. Una figura no es simplemente «buena» en abstracto: es adecuada o inadecuada para una pregunta, audiencia, tipo de datos y decisión determinados. La sección siguiente examina la idea de que la escala y el encuadre forman parte de la afirmación.

5. La escala y el encuadre forman parte de la afirmación

Los ejes, límites, anchos de intervalo, agregación y normalización determinan lo que un observador puede ver. Modificarlos puede ser legítimo, pero modifica la representación. La exigencia epistémica no consiste en congelar una vista universal, sino en declarar las transformaciones que afecten materialmente a la interpretación y examinar alternativas plausibles cuando la conclusión dependa de ellas.

En qué punto estamos

Los ejes, límites, anchos de intervalo, agregación y normalización determinan lo que un observador puede ver. Modificarlos puede ser legítimo, pero modifica la representación. La exigencia epistémica no consiste en congelar una vista universal, sino en declarar las transformaciones que afecten materialmente a la interpretación y examinar alternativas plausibles cuando la conclusión dependa de ellas. La sección siguiente examina la idea de que el color no es una etiqueta neutral.

6. El color no es una etiqueta neutral

El color puede separar clases y añadir dimensiones, pero el tono no posee una ordenación cuantitativa universal. Wong advirtió específicamente que los esquemas arcoíris son problemáticos para valores ordenados [6]. Crameri, Shephard y Heron mostraron que gradientes de color no uniformes pueden distorsionar la estructura aparente y resultar inaccesibles para lectores con deficiencias de visión cromática [7]. Las escalas modernas diseñadas perceptivamente mejoran esta situación, pero ninguna paleta elimina la necesidad de ajustar la codificación a la tarea.

En qué punto estamos

El color puede separar clases y añadir dimensiones, pero el tono no posee una ordenación cuantitativa universal. Crameri, Shephard y Heron mostraron que gradientes de color no uniformes pueden distorsionar la estructura aparente y resultar inaccesibles para lectores con deficiencias de visión cromática [7]. La representación tampoco es neutral. La sección siguiente examina la idea de que una imagen es datos más una cadena de renderizado.

7. Una imagen es datos más una cadena de renderizado

Una imagen microscópica no es idéntica a la muestra. Entre el sistema físico y los píxeles publicados intervienen parámetros de adquisición, reconstrucción, escalado de intensidad, filtrado y conversión de visualización. Rossner y Yamada subrayaron que la manipulación digital puede pasar de un ajuste legítimo a una alteración inapropiada de los datos originales [8]. La distinción clave no es «procesado frente a no procesado», sino si el procesamiento está científicamente justificado, se aplica de manera consistente cuando corresponde, se conserva y se declara.

En qué punto estamos

Una imagen microscópica no es idéntica a la muestra. Rossner y Yamada subrayaron que la manipulación digital puede pasar de un ajuste legítimo a una alteración inapropiada de los datos originales [8]. Entre el sistema físico y los píxeles publicados intervienen parámetros de adquisición, reconstrucción, escalado de intensidad, filtrado y conversión de visualización. La sección siguiente examina la idea de que el procesamiento puede revelar y ocultar.

8. El procesamiento puede revelar y ocultar

Las listas de comprobación comunitarias para microscopía, publicadas en línea en 2023 y editadas en Nature Methods en 2024, subrayan que los flujos de procesamiento y análisis de imagen transforman matemáticamente los datos de imagen y deben documentarse de forma transparente [9]. Los ajustes del rango de intensidad, el contraste no lineal, el filtrado, la restauración, la segmentación o los pasos de aprendizaje automático pueden mejorar la interpretabilidad, pero también eliminar información o generar apariencias engañosas. Por ello importan los datos brutos o mínimamente procesados y la procedencia del flujo de trabajo.

En qué punto estamos

Las listas de comprobación comunitarias para microscopía, publicadas en línea en 2023 y editadas en Nature Methods en 2024, subrayan que los flujos de procesamiento y análisis de imagen transforman matemáticamente los datos de imagen y deben documentarse de forma transparente [9]. La sección siguiente examina la idea de que un patrón visual todavía no es una explicación.

9. Un patrón visual todavía no es una explicación

Un cúmulo, gradiente o frontera aparente en un gráfico puede ser una pista. Por sí solo no identifica un mecanismo ni una causa. La representación puede reflejar el sistema subyacente, el preprocesamiento, el muestreo, la proyección, el suavizado o la propia codificación. La visualización genera y somete a prueba preguntas; la interpretación causal sigue necesitando la maquinaria probatoria desarrollada en KNO-014 y KNO-016.

En qué punto estamos

Un cúmulo, gradiente o frontera aparente en un gráfico puede ser una pista. Por sí solo no identifica un mecanismo ni una causa. La representación puede reflejar el sistema subyacente, el preprocesamiento, el muestreo, la proyección, el suavizado o la propia codificación. La sección siguiente examina la idea de que varias vistas funcionan como prueba de esfuerzo.

10. Varias vistas funcionan como prueba de esfuerzo

Como toda codificación privilegia ciertas relaciones, comparar vistas puede revelar dependencia de elecciones arbitrarias. Un diagrama de dispersión, un gráfico de residuos, un histograma y un resumen tabular pueden interrogar aspectos distintos del mismo conjunto de datos [3,4]. Las múltiples vistas no crean evidencia independiente si derivan de las mismas observaciones; esta última afirmación es una síntesis declarada del artículo y no una conclusión establecida directamente por esas fuentes.

En qué punto estamos

Como toda codificación privilegia ciertas relaciones, comparar vistas puede revelar dependencia de elecciones arbitrarias. Las múltiples vistas no crean evidencia independiente si derivan de las mismas observaciones; esta última afirmación es una síntesis declarada del artículo y no una conclusión establecida directamente por esas fuentes. La sección siguiente examina la idea de que claridad no equivale a verdad.

11. Claridad no equivale a verdad

Una figura clara puede presentar de forma magnífica un modelo incorrecto. Una figura saturada puede contener evidencia válida mal comunicada. La eficacia perceptiva, la validez estadística y la interpretación científica son dimensiones diferentes. Una buena visualización las alinea sin tratar ninguna como sustituto de las demás.

En qué punto estamos

Una figura clara puede presentar de forma magnífica un modelo incorrecto. Una figura saturada puede contener evidencia válida mal comunicada. La eficacia perceptiva, la validez estadística y la interpretación científica son dimensiones diferentes. Una buena visualización las alinea sin tratar ninguna como sustituto de las demás. La sección siguiente pregunta qué hace que una figura sea auditable y no solo persuasiva.

12. La figura auditable

Una visualización científica auditable permite identificar los datos de origen, las transformaciones, las codificaciones visuales y las omisiones importantes. En imágenes científicas, esto incluye la procedencia de adquisición y procesamiento; en gráficos, incluye escalas, agregación e incertidumbre cuando sea pertinente [8,9]. La reproducibilidad del flujo de renderizado no demuestra que la afirmación científica sea verdadera, pero la opacidad del flujo dificulta evaluarla.

En qué punto estamos

Una visualización científica auditable permite identificar los datos de origen, las transformaciones, las codificaciones visuales y las omisiones importantes. La reproducibilidad del flujo de renderizado no demuestra que la afirmación científica sea verdadera, pero la opacidad del flujo dificulta evaluarla. Esto delimita el alcance del artículo y prepara la síntesis final sin ampliar la afirmación más allá de lo establecido en las secciones anteriores.

Afirmaciones

- C01. Las representaciones diagramáticas y enunciativas pueden ser informacionalmente equivalentes y, aun así, diferir en eficiencia computacional para tareas concretas. [1]
- C02. Las codificaciones gráficas difieren en precisión de decodificación perceptiva. [2]
- C03. Estadísticas resumen similares pueden coexistir con estructuras de datos sustancialmente diferentes visibles en gráficos. [3]
- C04. El diseño de una visualización debe evaluarse en relación con datos, tarea, codificación y nivel de validación. [4]
- C05. Las propiedades perceptivas limitan el diseño eficaz de representaciones. [5]
- C06. El tono es un canal problemático como única codificación de magnitud cuantitativa ordenada y los mapas arcoíris pueden inducir a error. [6,7]
- C07. La accesibilidad y uniformidad perceptiva de los mapas de color son cuestiones de comunicación científica, no solo estéticas. [7]
- C08. La manipulación digital de imágenes científicas puede alterar su significado probatorio y debe controlarse. [8]
- C09. Los flujos de procesamiento y análisis de imágenes transforman datos numéricos y requieren información transparente. [9]
- C10. Un procesamiento de imagen legítimo no convierte la imagen mostrada en algo idéntico a la adquisición sin procesar. [8,9]
- C11. Varias vistas pueden exponer estructura o supuestos ocultos por una sola representación, sin convertirse por ello en evidencia independiente. **Síntesis declarada; las fuentes 3 y 4 aportan contexto, no apoyo directo para la cláusula de independencia.**
- C12. Un renderizado reproducible mejora la auditabilidad, pero no establece por sí solo la verdad de la interpretación científica. **Síntesis declarada; las fuentes 4, 8 y 9 aportan contexto sobre validación, integridad e información, no apoyo directo para la cláusula sobre verdad.**

Nota de frontera

KNO-018 trata la representación visual como método de conocimiento. No sustituye a KNO-006 sobre cartografía, KNO-013 sobre modelos, KNO-014 sobre identificación causal, KNO-015 sobre probabilidad/incertidumbre ni KNO-019 sobre tablas de datos y estructuras de registro. Tampoco afirma que exista una gramática visual universalmente óptima.

Lo que sabemos y cómo lo sabemos

Las representaciones visuales transforman datos en codificaciones espaciales y perceptivas. Esas decisiones afectan a qué patrones resultan fáciles de detectar, con qué precisión se decodifican las cantidades y qué aspectos de adquisición o procesamiento quedan visibles para el lector.

Lo que todavía no sabemos

No existe una codificación visual óptima para toda tarea, audiencia o conjunto de datos, y un renderizado atractivo o reproducible no establece la corrección de la interpretación científica. Puede quedar estructura importante fuera de la vista elegida.

Qué podría hacernos cambiar de opinión

La interpretación debe cambiar si escalas, codificaciones, procesamientos o vistas alternativas eliminan el patrón aparente, si la procedencia revela transformaciones no declaradas o si una característica visual no puede reproducirse desde los datos y el flujo indicados.

Si solo recuerdas una cosa

Una figura científica funciona como evidencia cuando su transformación desde los datos de origen es apropiada, inspeccionable y reconstruible.

Artículos relacionados

- KNO-006 · Maps and Spatial Representation
- KNO-013 · Models, Maps and Reality
- KNO-014 · Causality and Correlation
- KNO-019 · Data, Tables and Records

Referencias

- [1] Larkin, Jill H.; Simon, Herbert A. (1987). "Why a Diagram is (Sometimes) Worth Ten Thousand Words." *Cognitive Science* 11(1):65–100. DOI: 10.1111/j.1551-6708.1987.tb00863.x.
- [2] Cleveland, William S.; McGill, Robert (1984). "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." *Journal of the American Statistical Association* 79(387):531–554. DOI: 10.1080/01621459.1984.10478080.
- [3] Anscombe, F. J. (1973). "Graphs in Statistical Analysis." *The American Statistician* 27(1):17–21. DOI: 10.1080/00031305.1973.10478966.
- [4] Munzner, Tamara (2014). *Visualization Analysis and Design*. CRC Press. ISBN: 978-1-4665-0893-4.
- [5] Ware, Colin (2019). *Information Visualization: Perception for Design*. 4th ed. Morgan Kaufmann, sello de Elsevier. ISBN: 978-0-12-812875-6.
- [6] Wong, Bang (2010). "Color coding." *Nature Methods* 7:573. DOI: 10.1038/nmeth0810-573.
- [7] Crameri, Fabio; Shephard, Grace E.; Heron, Philip J. (2020). "The misuse of colour in science communication." *Nature Communications* 11:5444. DOI: 10.1038/s41467-020-19160-7.
- [8] Rossner, Mike; Yamada, Kenneth M. (2004). "What's in a picture? The temptation of image manipulation." *Journal of Cell Biology* 166(1):11–15. DOI: 10.1083/jcb.200406019.
- [9] Schmied, Christopher; Nelson, Michael S.; Avilov, Sergiy; et al. (2024; publicado en línea en 2023). "Community-developed checklists for publishing images and image analyses." *Nature Methods* 21(2):170–181. DOI: 10.1038/s41592-023-01987-9.

KNO-019 · Datos, tablas y registros

KNO · Formas de conocer

Juan Segurola

v0.7 · ES

Todos los derechos reservados

Resumen del artículo

El trabajo científico deja tablas, hojas de cálculo, conjuntos de datos y archivos, pero la supervivencia de un fichero no indica por sí sola qué podrán aprender de él investigadores posteriores. ¿Qué debe conservar un registro para que, años después, alguien pueda interpretar responsablemente una observación, distinguir ausencia de no detección, identificar la versión analizada y reconstruir las transformaciones entre adquisición y publicación? Este artículo examina esas preguntas a través de estructura tabular, prácticas de hoja de cálculo, metadatos, procedencia, versionado, identificadores persistentes y la categoría ambigua de «datos brutos». Pregunta dónde termina el simple almacenamiento y comienza un registro científico auditable, cuánto contexto debe viajar con los valores numéricos y qué partes de la historia de producción de los datos deben seguir visibles para reanálisis, corrección y reutilización a largo plazo.

En 60 segundos

Los datos se convierten en evidencia duradera solo cuando su estructura y su historia siguen siendo interpretables. Las tablas facilitan algunas comparaciones, pero la comodidad de las hojas de cálculo también puede ocultar unidades, tipos, fórmulas o convenciones de valores ausentes. Los metadatos conservan el contexto necesario para entender qué significan los valores. La procedencia registra cómo se produjeron o transformaron los datos. El versionado importa porque citar «el conjunto de datos» es ambiguo si el objeto cambia, mientras que los identificadores persistentes estabilizan la referencia sin sustituir la documentación. Los valores ausentes deben representarse explícitamente porque ausencia, no detección y no aplicabilidad no son el mismo estado. Los «datos brutos» tampoco son necesariamente intactos: instrumentos y sistemas de adquisición pueden haber aplicado calibración o procesamiento. Preservar tiene éxito solo si futuros usuarios pueden reconstruir suficientemente la cadena desde observación hasta análisis. Un buen registro no es solo un archivo que sobrevive, sino uno cuyo significado, linaje y versión siguen siendo auditables.

Por qué importa esta pregunta

La ciencia moderna depende de registros que sobreviven al momento de la observación. Las mediciones se copian en tablas, se transforman con software, se combinan con metadatos, se depositan en repositorios y se citan en trabajos posteriores. Cada paso puede conservar información o destruir contexto silenciosamente. Un conjunto de datos puede estar numéricamente intacto y resultar inutilizable si faltan definiciones de variables, unidades, exclusiones o historia de procesamiento. El problema importa para reproducibilidad, reanálisis, corrección de errores y preservación a largo plazo. Entender los registros como objetos estructurados, versionados y con procedencia permite distinguir «los datos siguen existiendo» de la afirmación más fuerte de que otra persona todavía puede comprender qué significan y cómo llegaron a ser evidencia.

1. Un registro no es el acontecimiento

Un registro es una representación que deja un acontecimiento, una observación, una medición o una decisión. Puede conservar números, texto, imágenes, marcas temporales o categorías, pero no es idéntico a aquello que los produjo. **C01. Los registros amplían la memoria al hacer persistentes las observaciones, pero la persistencia no garantiza la interpretabilidad. Síntesis declarada:** [3] y [5] respaldan la procedencia, los metadatos y la reconstrucción; no establecen directamente la cláusula sobre interpretabilidad.

Esta distinción importa porque los usuarios posteriores suelen encontrarse con el registro cuando el contexto original ya ha desaparecido. La tarea epistémica no consiste solo en almacenar contenido, sino en conservar suficiente información sobre su producción para que otra persona pueda comprender qué significa.

En qué punto estamos

Un registro representa un acontecimiento u observación; no es el acontecimiento mismo, de modo que preservar siempre depende de qué se codificó y qué quedó fuera. La afirmación es conceptual y no niega que los registros puedan ser evidencia fiable. La siguiente pregunta es estructural: ¿qué organización necesitan los datos para inspeccionar de forma consistente las relaciones entre observaciones?

2. Los datos necesitan estructura

Una colección de valores resulta más fácil de inspeccionar y analizar cuando su estructura es explícita. La formulación de datos ordenados de Wickham es una convención influyente: las variables ocupan columnas, las observaciones ocupan filas y los distintos tipos de unidades observacionales ocupan tablas distintas. [1]

C02. Una estructura tabular coherente puede reducir la ambigüedad y facilitar la transformación, el modelado y la visualización, pero es una convención para organizar datos, no una ontología universal del mundo. [1]

Una fila no es de manera natural una «observación» hasta que alguien define la unidad observacional. El mismo acontecimiento puede representarse al nivel de una persona, una visita, una lectura instrumental, una imagen, un ensayo o una transacción. La tabla incorpora, por tanto, una decisión de modelado.

En qué punto estamos

Los datos estructurados reducen ambigüedad al hacer explícitas variables, observaciones y valores, aunque ningún diseño tabular sea neutral respecto de operaciones posteriores. La ganancia epistémica es consistencia e inspeccionabilidad, no corrección automática. Una vez explícita la estructura, la siguiente pregunta es qué revelan bien las tablas y qué pueden ocultar al comprimir relaciones en filas y columnas.

3. Las tablas muestran unas relaciones y ocultan otras

Las tablas son poderosas porque alinean valores comparables. Permiten ordenar, filtrar, unir, agrupar y calcular. Sin embargo, una tabla puede ocultar la historia que produjo sus celdas. Una columna final puede contener un valor generado mediante varias transformaciones cuyos pasos intermedios no están presentes.

C03. La organización tabular facilita la comparación y el cálculo, pero la tabla por sí sola no tiene por qué conservar la procedencia de cada valor. Síntesis declarada: [2] respalda la organización tabular disciplinada y [3] respalda la procedencia; ninguna de las dos fuentes afirma directamente la cláusula de no conservación.

Por eso, una tabla limpia y un registro auditable no son lo mismo. Una buena organización ayuda al análisis; la procedencia ayuda a reconstruirlo.

En qué punto estamos

Las tablas exponen comparaciones, uniones y patrones con eficiencia, pero suprimen otras relaciones que pueden requerir representaciones distintas. Su utilidad depende de ajustar la estructura a la pregunta. Ese límite conduce a un peligro práctico: cuando los datos tabulares se manejan en hojas de cálculo flexibles, ¿qué comodidades pueden debilitar silenciosamente la integridad y auditabilidad del registro?

4. La comodidad de las hojas de cálculo puede crear fragilidad epistémica

Las hojas de cálculo son útiles porque combinan entrada de datos, cálculo y presentación. Esa misma flexibilidad también puede mezclar datos con formato, comentarios, varias variables en una sola celda, cálculos incrustados o códigos inconsistentes para valores ausentes. Broman y Woo recomiendan una organización rectangular simple, una sola pieza de información por celda, convenciones coherentes para valores ausentes y separación entre datos originales y cálculos. [2]

C04. Las prácticas de hoja de cálculo que mezclan observaciones originales, formato y cálculos derivados pueden hacer más propensas a error la interpretación y el análisis posteriores. [2]

No es un argumento contra las hojas de cálculo. Es un argumento a favor de tratar la estructura del registro como parte del método.

En qué punto estamos

Las hojas de cálculo son entornos de trabajo potentes, pero tipos mezclados, fórmulas ocultas, formato manual y celdas sobrecargadas pueden volver los datos frágiles o ambiguos. El riesgo no es que sean inválidas por naturaleza, sino que la comodidad oculte estructura. La siguiente pregunta es qué información debe acompañar a los valores para que futuros usuarios puedan interpretarlos correctamente.

5. Los metadatos conservan el significado que rodea a los valores

Los metadatos describen los datos. Pueden identificar variables, unidades, instrumentos, fechas, lugares, creadores, métodos, software, condiciones de acceso o vínculos con registros relacionados. El Research Data Framework de NIST trata los metadatos como esenciales para el uso, reutilización y preservación eficaces a lo largo del tiempo. [5]

C05. Los valores se vuelven más reutilizables cuando los metadatos conservan las definiciones y condiciones necesarias para interpretarlos. [5]

Una columna etiquetada `temperature` es incompleta si la escala, el sensor, el lugar, el momento y el procesamiento son relevantes para la pregunta. Los metadatos no eliminan mágicamente la incertidumbre; conservan el contexto necesario para razonar sobre ella.

En qué punto estamos

Los metadatos conservan el contexto alrededor de los valores: definiciones de variables, unidades, métodos, identificadores y demás información necesaria para interpretar el registro. Unos metadatos completos no garantizan ciencia válida, pero la falta de contexto puede inutilizar observaciones válidas. La siguiente pregunta es histórica: ¿cómo conservar no solo qué son los valores, sino cómo llegaron a existir?

6. La procedencia registra cómo llegaron a existir los datos

La familia W3C PROV define la procedencia en términos de entidades, actividades y agentes implicados en la producción de un dato u otro objeto. Está concebida para permitir descripciones interoperables de los orígenes y las transformaciones. [3]

C06. Los registros de procedencia pueden documentar de dónde proceden los datos, quién o qué actuó sobre ellos y qué procesos contribuyeron al objeto resultante. [3]

La procedencia es, por tanto, distinta de una etiqueta descriptiva. Un conjunto de datos puede tener nombres de variables excelentes y, aun así, no mostrar qué archivo fuente, ejecución instrumental, script o transformación produjo la versión analizada.

En qué punto estamos

La procedencia registra el linaje de los datos a través de adquisición, transformación y uso, permitiendo inspeccionar cómo se produjo un objeto. Mejora la reconstructibilidad sin demostrar que cada transformación haya sido apropiada. Si el linaje importa, aparece otro problema: cuando un conjunto de datos cambia con el tiempo, ¿qué estado exacto de ese objeto evolutivo identifica una cita posterior?

7. El versionado cambia aquello a lo que apunta una cita

Los datos pueden cambiar después de su creación inicial: se corrigen errores, se añaden registros, se revisan clasificaciones y se actualizan los procesos de tratamiento. Si una afirmación depende de un estado concreto de un conjunto de datos, citar solo el nombre general del conjunto puede ser insuficiente.

La Joint Declaration of Data Citation Principles subraya la especificidad y verificabilidad, incluida información suficiente para identificar una versión, un intervalo temporal o una porción granular concretos de los datos. [4]

C07. Cuando los datos cambian con el tiempo, la información de versión puede ser necesaria para determinar qué datos exactos sustentaron una afirmación. [4][5]

La afirmación de conocimiento queda vinculada no solo a una familia de datos, sino a una manifestación identificable de ella.

En qué punto estamos

El versionado distingue estados sucesivos de un objeto de datos, permitiendo que una referencia señale lo realmente analizado y no un objetivo móvil inespecífico. Una etiqueta de versión solo ayuda si esa versión se conserva e identifica. La siguiente pregunta es cómo contribuyen los identificadores persistentes y qué problemas de documentación no pueden resolver por sí solos.

8. Los identificadores persistentes resuelven un problema, no todos

Los identificadores persistentes ayudan a mantener estables las referencias cuando los objetos cambian de ubicación. Los principios de citación de datos recomiendan una identificación globalmente única, persistente y procesable por máquinas. [4]

C08. Un identificador persistente mejora la estabilidad referencial, pero por sí solo no establece la calidad de los datos, la corrección de la interpretación ni la relevancia probatoria. [4]

Un DOI puede decirnos qué objeto se citó. No puede decirnos si las mediciones eran válidas o si el análisis respondía a la pregunta pretendida. Identificación y evaluación son tareas epistémicas distintas.

En qué punto estamos

Los identificadores persistentes hacen las referencias más estables y resolubles por máquinas, pero no aportan por sí mismos metadatos, procedencia ni semántica de versiones ausentes. Resuelven identificación, no todo el problema de interpretabilidad. La siguiente pregunta vuelve al interior de la tabla: cuando una celda carece de valor ordinario, ¿cómo distinguir ausencia, cero, no detección o no aplicabilidad?

9. Los valores ausentes también son registros

La ausencia en una tabla puede significar muchas cosas: no medido, no aplicable, por debajo del límite de detección, perdido, rechazado, imposible, corrupto o realmente cero. Tratar estos estados como intercambiables puede cambiar el análisis.

C09. La ausencia de datos requiere representación explícita porque un campo vacío, un cero y un valor no observado pueden tener significados distintos. [2]

El esquema de codificación exacto depende del dominio. El punto metodológico general es que un registro debe distinguir estados diferentes cuando esas diferencias importan para la interpretación.

En qué punto estamos

Los valores ausentes son estados que requieren representación explícita, porque razones distintas para la ausencia pueden implicar análisis e interpretaciones diferentes. Codificar bien la ausencia no recupera información nunca observada, pero evita colapsar varios estados en uno. La siguiente pregunta es si la etiqueta habitual «datos brutos» identifica de manera fiable un punto de partida realmente intacto.

10. Los datos «brutos» no siempre son datos intactos

La expresión «datos brutos» suele sugerir una capa prístina anterior a la interpretación. En la práctica, los instrumentos pueden aplicar ya calibración, filtrado, compresión, umbrales o conversiones propietarias antes de que el investigador reciba un archivo. El RDaF de NIST trata explícitamente los instrumentos, las operaciones de procesamiento, los parámetros, los flujos de trabajo y la procedencia como partes relevantes de la gestión de datos. [5]

C10. El registro digital más temprano disponible puede contener ya decisiones de procesamiento, por lo que «bruto» no debe suponerse equivalente a «sin transformar». Síntesis declarada: [5] documenta instrumentos, operaciones de procesamiento, parámetros, flujos de trabajo y procedencia; no afirma directamente que el archivo disponible más temprano esté necesariamente transformado.

Lo importante es documentar el estado más temprano que se conserva y las transformaciones entre estados.

En qué punto estamos

Los llamados datos brutos pueden incorporar ya calibración, filtrado, digitalización u otras transformaciones realizadas por instrumentos y sistemas de adquisición. «Bruto» es por tanto una etiqueta relativa al flujo de trabajo, no garantía de realidad intacta. Esto vuelve más importante la procedencia. La siguiente pregunta es qué ocurre cuando los archivos se preservan pero desaparece el conocimiento contextual necesario para interpretarlos.

11. Los datos pueden conservarse y, sin embargo, volverse inutilizables

Un archivo puede permanecer intacto byte a byte mientras desaparecen el software, el esquema, el diccionario de datos, las unidades o la procedencia necesarios para interpretarlo. NIST señala que unos metadatos deficientes pueden volver inutilizables conjuntos de datos importantes cuando sus creadores ya no están disponibles. [5]

C11. La preservación de los bytes no equivale a la preservación del significado ni de la reutilizabilidad. [5]

Por eso, los registros a largo plazo dependen de documentación, formatos accesibles, identificadores y metadatos contextuales, no solo de almacenamiento.

En qué punto estamos

El almacenamiento a largo plazo puede conservar bytes mientras se pierde el conocimiento necesario para interpretar variables, decisiones de procesamiento o contexto instrumental. Preservar es necesario pero no suficiente para reutilizar. El límite epistémico es semántico: supervivencia del archivo no implica supervivencia del significado. La siguiente pregunta es cuánto debe reconstruirse de la cadena observación-análisis para sostener reproducibilidad.

12. Los registros favorecen la reproducibilidad solo cuando la cadena puede reconstruirse

La reproducibilidad exige más que conservar los números finales. El tratamiento de la reproducibilidad por las National Academies subraya el acceso a los datos y a la información computacional necesaria para volver a calcular resultados, mientras que NIST vincula la reproducibilidad con metadatos, código, métodos e instrumentos documentados. [5][6]

C12. Los registros contribuyen a la reproducibilidad cuando conservan suficiente información sobre datos, metadatos, código, métodos y procedencia para reconstruir el resultado relevante. [5][6]

Esto no promete que todo estudio pueda recrearse exactamente. Algunas observaciones son únicas, destructivas o históricamente irrepetibles. La función del registro es conservar, hasta donde permita el contexto de investigación, el trayecto reconstruible desde la observación hasta la afirmación.

En qué punto estamos

Los registros apoyan reproducibilidad cuando otro investigador puede reconstruir la cadena relevante de datos, código, transformaciones, versiones y contexto para repetir el recorrido computacional. La reconstructibilidad no garantiza una interpretación científica idéntica. La siguiente pregunta es práctica: ¿qué auditoría mínima debe realizarse antes de tratar un conjunto de datos o una tabla como registro evidencial duradero?

13. Una auditoría práctica del registro

Para un conjunto de datos importante, pregunte:

1. ¿Qué representa una fila o registro?
2. ¿Qué significa cada variable y cuáles son sus unidades?
3. ¿Cómo se distinguen los valores ausentes, censurados y no aplicables?
4. ¿Qué registro es original, cuál está depurado y cuál es derivado?
5. ¿Qué transformaciones los conectan?
6. ¿Qué instrumento, método, software y parámetros produjeron los valores?
7. ¿Puede identificarse la versión exacta usada para una afirmación?

8. ¿Puede otra persona interpretar los datos sin depender de la memoria de quien los creó?

Esta lista es una síntesis editorial para este artículo, no un estándar universal de gestión de datos.

En qué punto estamos

Una auditoría práctica comprueba estructura, unidades, ausencia, metadatos, procedencia, versiones, identificadores y vínculos entre salidas transformadas y entradas. Superarla mejora la inspeccionabilidad, pero no demuestra que las mediciones subyacentes sean correctas. La siguiente pregunta es de frontera: ¿qué temas KNO vecinos poseen visualización, síntesis de evidencia y reproducibilidad más allá del problema más estrecho de los registros?

14. Frontera de este artículo

KNO-019 trata del papel epistémico de los registros y los datos estructurados. No intenta prescribir un esquema universal de bases de datos, una política de repositorio ni un flujo estadístico. KNO-018 trató las representaciones visuales; KNO-022 tratará la computación como memoria externa; KNO-024 tratará la pérdida y el sesgo de preservación. Aquí la pregunta más estrecha es: **¿qué debe sobrevivir en un registro para que el razonamiento posterior siga siendo posible?**

En qué punto estamos

KNO-019 posee el problema epistémico de registros estructurados, metadatos, procedencia y objetos de datos versionados; no sustituye a KNO-018 sobre visualización ni a KNO-016 sobre síntesis de evidencia, y solo prepara KNO-021 sobre reproducibilidad abierta. Fijada esa frontera, el artículo puede resumir qué preservan los registros, qué no pueden preservar por sí solos y qué cambiaría nuestra confianza en ellos.

Lo que sabemos y cómo lo sabemos

Sabemos que una organización tabular explícita puede facilitar el análisis de variables y observaciones [1], que las convenciones de hoja de cálculo pueden reducir ambigüedades evitables y la mezcla de contenido original y derivado [2], que la procedencia puede representar entidades, actividades y agentes implicados en la producción de datos [3], y que la citación de datos exige suficiente especificidad para identificar el objeto o la versión citados [4]. NIST RDaF Version 2.0 trata explícitamente los metadatos, la procedencia, el versionado, los instrumentos, las operaciones de procesamiento y la reproducibilidad como partes de la gestión de datos de investigación [5]. Las National Academies vinculan asimismo la reproducibilidad con el acceso a los datos y a la información computacional necesarios para volver a calcular resultados [6].

Lo que todavía no sabemos

Ninguna estructura universal de registro garantiza la interpretabilidad futura. Distintas disciplinas requieren diferentes unidades, metadatos, granularidad de procedencia, periodos de conservación y formas de representar datos ausentes. Tampoco la documentación por sí sola establece que las mediciones sean válidas o que el análisis responda a la pregunta científica correcta.

Qué podría hacernos cambiar de opinión

Las conclusiones metodológicas de este artículo tendrían que revisarse si apareciera evidencia de que estas distinciones fallan de forma sistemática en dominios importantes, si modelos de procedencia o versionado más eficaces sustituyeran las prácticas actuales, o si futuras infraestructuras hicieran recuperables por otros medios algunos registros contextuales que hoy son necesarios. Los cambios de herramienta no eliminarían el requisito de fondo: que un lector pueda identificar e interpretar el objeto probatorio realmente utilizado.

Si solo recuerdas una cosa

Los datos conservan conocimiento solo cuando el registro preserva suficiente significado, procedencia y contexto de versión para que los valores puedan interpretarse.

Artículos relacionados

- KNO-016 · *Cómo se construye la evidencia: integrando las herramientas*
- KNO-018 · *Imágenes, diagramas y visualización*
- KNO-022 · *La computación como memoria externa*
- KNO-024 · *Conocimiento perdido y sesgo de preservación*

Referencias

[1] Hadley Wickham. “Tidy Data.” *Journal of Statistical Software* 59(10), 2014. DOI: 10.18637/jss.v059.i10.

[2] Karl W. Broman; Kara H. Woo. “Data Organization in Spreadsheets.” *The American Statistician* 72(1):2–10, 2018. DOI: 10.1080/00031305.2017.1375989.

[3] World Wide Web Consortium (W3C). *PROV-Overview: An Overview of the PROV Family of Documents*. W3C Working Group Note, 30 de abril de 2013. URI estable de la nota del W3C: <https://www.w3.org/TR/2013/NOTE-prov-overview-20130430/>

[4] Data Citation Synthesis Group. *Joint Declaration of Data Citation Principles*. FORCE11, 2014. Editora: Maryann E. Martone. DOI: 10.25490/a97f-egyk.

[5] Robert Hanisch; Debra L. Kaiser; Alda Yuan; Andrea Medina-Smith; Bonnie C Carroll; Eva Campo. *NIST Research Data Framework (RDaF): Version 2.0*. NIST Special Publication 1500-18r2, 7 de febrero de 2024. DOI: 10.6028/NIST.SP.1500-18r2.

[6] National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press, 2019. DOI: 10.17226/25303.

KNO-020 · Revisión, consenso y desacuerdo

KNO · Formas de conocer

Juan Seguro

v0.6 · ES

Todos los derechos reservados

Resumen del artículo

La ciencia utiliza varios procedimientos sociales para exponer las afirmaciones a crítica, pero esos procedimientos no desempeñan la misma función. Un manuscrito puede revisarse antes de publicarse, un resultado publicado puede afrontar replicación o corrección, y una comunidad puede describir más tarde una posición como consenso mientras siguen abiertas cuestiones importantes. El desacuerdo puede surgir de la evidencia, los métodos, los modelos, los supuestos, la incertidumbre o los valores, y esas fuentes no requieren necesariamente el mismo remedio. Este artículo pregunta cómo deben interpretarse la revisión, el consenso y el desacuerdo como evidencia sobre la evaluación científica. Separa la evidencia de primer orden de los juicios sobre ella, sigue qué cambia con la publicación y pregunta qué hace informativos al acuerdo y al desacuerdo.

En 60 segundos

La revisión por pares, la publicación, el consenso y el desacuerdo son etapas o señales dentro de la evaluación científica, no sustitutos de la evidencia subyacente. Los revisores pueden descubrir debilidades antes de publicar, aunque revisores distintos pueden destacar riesgos diferentes de manera razonable. La publicación amplía el escrutinio en lugar de terminarlo. El consenso resulta más informativo cuando registra convergencia entre evidencia, métodos y examen crítico, sin dejar de ser revisable. Una posición minoritaria no queda descalificada por ser minoritaria, pero tampoco adquiere por ello apoyo evidencial equivalente. El desacuerdo productivo identifica dónde divergen las interpretaciones y qué observación, análisis o replicación podría discriminarlas. Las correcciones y retractaciones pueden revelar debilidades, pero también mostrar que el registro académico sigue abierto a reparación. Para el lector, la tarea práctica es preguntar qué se revisó, qué evidencia sostiene la afirmación, dónde reside el desacuerdo y qué podría cambiar el balance evidencial en la práctica.

Por qué importa esta pregunta

La discusión pública comprime a menudo la evaluación científica en eslóganes: «revisado por pares» se trata como un sello de verdad, «consenso científico» como una votación y «los científicos discrepan» como prueba de que el conocimiento se ha derrumbado. Esos atajos ocultan la diferencia entre la evidencia y los procedimientos sociales usados para someterla a escrutinio. La distinción importa cuando los artículos se contradicen, los comités de expertos emiten evaluaciones, una interpretación minoritaria desafía la dominante o una corrección modifica el registro publicado. Entender qué pueden y qué no pueden establecer revisión y consenso permite juzgar controversias sin venerar la autoridad ni confundir desacuerdo con parálisis epistémica.

1. La revisión por pares es un filtro, no una máquina de verdad

La revisión de una revista pide a lectores cualificados que examinen un manuscrito antes de una decisión editorial. El proceso puede identificar razonamientos poco claros, controles ausentes, afirmaciones sin apoyo, problemas metodológicos o bibliografía relevante que los autores hayan pasado por alto. COPE atribuye a los revisores un papel importante en la integridad del registro académico y exige revisiones objetivas, constructivas y adecuadamente ajustadas a la experiencia del revisor. [1]

Ese papel es importante, pero limitado. Un revisor suele ver un manuscrito, sus métodos declarados y la evidencia aportada por los autores. No repite todos los experimentos, no regenera de forma independiente todos los conjuntos de datos ni garantiza que los supuestos no declarados sean inocuos. La investigación empírica sobre revisión por pares también ha documentado variabilidad en los juicios de los revisores y preocupaciones sobre fiabilidad, sesgo y validez predictiva. [2][3]

Afirmación C01 - Síntesis declarada. La revisión por pares se entiende mejor como una oportunidad estructurada de crítica experta antes o alrededor de la publicación, no como un procedimiento que establezca la verdad. La primera parte está directamente respaldada por [1][2]; el contraste «no verdad» es una síntesis interpretativa a partir del alcance y las limitaciones documentados del proceso.

En qué punto estamos

La revisión por pares puede revelar debilidades antes de publicar, pero su alcance está limitado por lo que los revisores pueden inspeccionar y por la información que aportan los autores. El estado epistémico sigue siendo provisional: la revisión añade crítica estructurada sin certificar verdad. Aclarado ese límite, la siguiente pregunta es qué significa que revisores competentes discrepen entre sí.

2. El desacuerdo entre revisores no es por sí mismo un fallo

Dos revisores competentes pueden leer de forma distinta el mismo artículo. Pueden especializarse en métodos diferentes, detectar modos de fallo distintos o juzgar de forma diferente la importancia de las limitaciones. Ese desacuerdo puede ser útil porque revela lugares donde el manuscrito depende de supuestos que no eran evidentes para todos los lectores.

La práctica editorial no reduce esto a contar votos. *Nature Genetics* indica explícitamente a sus revisores que otros revisores pueden tener experiencia o puntos de vista distintos, y afirma que las decisiones editoriales no se toman contando votos ni puntuaciones numéricas. Los editores ponderan los argumentos planteados por revisores y autores. [4]

Afirmación C02. El desacuerdo entre revisores puede revelar diferencias de experiencia, supuestos y ponderación de la evidencia, y no simplemente que uno de ellos sea incompetente. [2][4]

En qué punto estamos

El desacuerdo entre revisores puede ser informativo porque experiencias y supuestos distintos exponen vulnerabilidades diferentes, y las decisiones editoriales no tienen por qué reducir esos juicios a una votación. La evidencia respalda el desacuerdo como señal diagnóstica, no como prueba de que un manuscrito sea correcto o incorrecto. La siguiente pregunta es qué cambia cuando una afirmación pasa de revisión a publicación.

3. La publicación es una etapa del escrutinio, no su final

Una vez publicado un artículo, el escrutinio se amplía. Otros investigadores pueden intentar reproducir cálculos, repetir experimentos, estudiar otras poblaciones, cuestionar mediciones o comparar explicaciones alternativas. Las Academias Nacionales distinguen la reproducibilidad usando los mismos datos y pasos computacionales de la replicabilidad entre estudios con datos nuevos, y subrayan que un fallo de replicación puede en algunos casos revelar limitaciones o abrir nuevas preguntas científicas. [5]

La publicación cambia, por tanto, el estatus social de una afirmación sin congelar su estatus epistémico. Un artículo puede corregirse, matizarse, contradecirse o retractarse posteriormente. El influyente ensayo de Ioannidis de 2005 formalizó cómo el sesgo, la baja potencia, la flexibilidad y las probabilidades previas pueden hacer poco fiables las afirmaciones publicadas en determinados contextos; el propio artículo recibió posteriormente una corrección en una tabla, recordatorio útil de que publicación y corrección posterior forman parte del mismo registro. [6][7]

Afirmación C03. La publicación amplía la comunidad capaz de poner a prueba una afirmación; no termina el proceso por el que esa afirmación es evaluada. [5][6]

En qué punto estamos

La publicación hace una afirmación más ampliamente inspeccionable y le da un estatus académico duradero, pero no congela su estatus epistémico. Replicación, reanálisis, corrección y contradicción aún pueden cambiar el balance de evidencia. Esto abre otra pregunta: si publicar no significa finalizar, ¿qué debe significar exactamente el consenso científico cuando una comunidad informa de amplio acuerdo?

4. Consenso no significa unanimidad

El consenso científico puede significar varias cosas y el término debe utilizarse con cuidado. En procesos formales de estudios de consenso, las Academias Nacionales describen sus informes de consenso como posiciones basadas en evidencia producidas por comités de expertos a partir de información reunida y deliberación, seguidas de revisión independiente. [8]

Eso es distinto de pedir a una sala que levante la mano. El valor epistémico procede de la evidencia y de los procedimientos utilizados para desafiar interpretaciones, no de la unanimidad como tal. Un consenso maduro puede seguir conteniendo incertidumbre, subpreguntas no resueltas e interpretaciones minoritarias legítimas.

Afirmación C04 - Síntesis declarada. En evaluación científica, el consenso es más informativo cuando registra juicio experto convergente apoyado en un cuerpo explícito de evidencia y deliberación crítica. [8] respalda directamente el modelo institucional; la generalización fuera de ese modelo es una síntesis.

En qué punto estamos

El consenso resulta más informativo cuando registra convergencia experta basada en evidencia bajo procedimientos que exponen el razonamiento a crítica; no equivale a unanimidad ni a contar cabezas. La conclusión está bien respaldada institucionalmente, aunque su generalización exige cuidado. La siguiente pregunta es por qué el acuerdo alcanzado por rutas parcialmente independientes pesa más que la repetición basada en dependencias compartidas.

5. El acuerdo resulta más informativo cuando las rutas hacia él son parcialmente independientes

Diez artículos que reutilizan el mismo conjunto de datos, supuesto o calibración no equivalen a diez líneas independientes de investigación. La convergencia informa más cuando diferentes mediciones, diseños, modelos o poblaciones pueden fallar de formas distintas. Es la misma razón por la que importan la replicación y la diversidad metodológica: debilidades compartidas pueden producir conclusiones compartidas.

Las Academias Nacionales señalan que la consistencia entre estudios independientes puede reforzar la confianza y, al mismo tiempo, advierten que la falta de replicación puede tener múltiples causas y debe interpretarse en contexto. [5]

Afirmación C05. El peso evidencial del acuerdo depende en parte de cuán independientes sean la evidencia y los métodos que contribuyen a él. Esta es una síntesis declarada basada en la lógica de replicación descrita en [5].

En qué punto estamos

El acuerdo gana fuerza evidencial cuando los estudios, métodos o medidas contribuyentes no comparten todos los mismos modos de fallo. La independencia rara vez es absoluta, por lo que la afirmación es comparativa y no binaria. Queda una dificultad práctica: incluso una mayoría fuerte puede equivocarse; ¿cómo tratar entonces una posición minoritaria sin descartarla ni concederle paridad automática?

6. Una posición minoritaria no es automáticamente errónea

La historia contiene casos en los que posiciones minoritarias se volvieron mayoritarias tras aparecer nuevas evidencias, mejores instrumentos o teorías mejoradas. Esa posibilidad importa porque la revisión por pares y el juicio experto pueden ser conservadores. Benda y Engels encontraron evidencia de que la revisión por pares puede funcionar razonablemente bien en promedio y, al mismo tiempo, identificaron tensión entre revisión e ideas innovadoras, con sesgos y preconcepciones como riesgos. [3]

Pero el error simétrico es igual de serio: la existencia de una posición minoritaria no demuestra que la mayoría esté mal respaldada. Un desacuerdo tiene fuerza evidencial solo en la medida en que la posición competidora aporte métodos, observaciones o argumentos capaces de cambiar el balance de evidencia.

Afirmación C06 - Síntesis declarada. El estatus minoritario no debe ni descalificar una afirmación ni otorgarle paridad evidencial. [3] respalda el riesgo de conservadurismo; la regla de paridad es una síntesis metodológica explícita.

En qué punto estamos

La condición minoritaria no es por sí misma un defecto evidencial y los sistemas de revisión pueden ser conservadores; pero disentir tampoco crea apoyo equivalente cuando la evidencia es desigual. La posición justificada depende de métodos, observaciones y argumentos, no del rango social. La siguiente pregunta es cómo localizar el desacuerdo para convertir interpretaciones enfrentadas en alternativas realmente contrastables.

7. El desacuerdo es más útil cuando puede localizarse

«Los científicos discrepan» suele ser demasiado grueso para informar. El desacuerdo puede referirse a las observaciones brutas, a cómo se midieron las variables, a qué casos se excluyeron, a qué modelo es apropiado, a cómo se propagó la incertidumbre, a si una interpretación causal está identificada o a cuánto peso asignar a evidencia competidora.

Localizar el desacuerdo importa porque desacuerdos distintos exigen pruebas distintas. Una disputa de calibración puede resolverse con mejor medición. Una disputa de modelos puede exigir una predicción discriminante. Una disputa causal puede requerir otro diseño. Un desacuerdo sobre valores o riesgo aceptable no puede resolverse fingiendo que es solo un problema de medición.

Afirmación C07 - Síntesis declarada. Los desacuerdos científicos más productivos especifican el nivel en que divergen las interpretaciones competidoras y qué observación o análisis podría discriminarlas. Esta es una síntesis metodológica de la lógica de revisión, replicación y contraste de modelos desarrollada en KNO-013, KNO-016 y el presente artículo.

En qué punto estamos

Un desacuerdo se vuelve científicamente útil cuando especificamos si afecta a observaciones, medida, elección de modelo, causalidad, incertidumbre o valores, porque fuentes distintas requieren pruebas distintas. Es una síntesis metodológica, no una taxonomía universal. Localizado el desacuerdo, la siguiente pregunta es si un consenso estable debe considerarse cerrado o simplemente bien respaldado por ahora.

8. El consenso puede ser estable sin ser final

Un consenso puede permanecer estable durante décadas porque pruebas sucesivas siguen respaldándolo. También puede cambiar con rapidez cuando un nuevo resultado elimina un supuesto central o un nuevo instrumento abre un régimen antes inaccesible. La estabilidad dice algo sobre la historia de desafíos superados, pero no crea inmunidad lógica frente a la revisión.

El tratamiento de reproducibilidad y replicabilidad de las Academias Nacionales permite explícitamente que una no replicación contribuya al descubrimiento en vez de tratar toda inconsistencia como automáticamente patológica. [5]

Afirmación C08. Un consenso científico puede estar bien respaldado y seguir siendo revisable en principio. [5][8]

En qué punto estamos

Un consenso estable puede reflejar supervivencia repetida ante la crítica sin adquirir inmunidad lógica frente a revisión. Su estabilidad es evidencia sobre una historia de desafíos superados, no una garantía sobre toda observación futura. Esa distinción conduce al propio registro académico: si las afirmaciones siguen siendo revisables tras publicar, ¿qué debemos aprender cuando aparece una corrección explícita?

9. Una corrección es evidencia de que el escrutinio continuó

Las correcciones y retractaciones se presentan a veces como prueba de que la ciencia «falló». También pueden mostrar que el registro permaneció abierto a reparación. Esto no significa que toda corrección sea trivial ni que la autocorrección institucional sea siempre rápida. Significa que la pregunta relevante es si los errores pueden detectarse, documentarse y propagarse a los trabajos posteriores.

La corrección de 2022 al artículo de Ioannidis de 2005 es un ejemplo sencillo: la revista conservó el artículo original, enlazó la corrección y especificó el error exacto de la tabla. [7]

Afirmación C09. Una corrección visible puede ser evidencia de control de calidad continuado, aunque la existencia de una corrección no mida por sí sola la fiabilidad del resto del registro. [7]

En qué punto estamos

Una corrección visible puede mostrar que el escrutinio continuó y que un error concreto quedó documentado en el registro público, aunque la existencia de una corrección no certifique la fiabilidad del resto. La ganancia epistémica es auditabilidad, no absolución. La siguiente pregunta es qué debe exponer una afirmación de revisión o consenso para poder auditarla en vez de confiar en su etiqueta.

10. Las afirmaciones de consenso necesitan una pista de auditoría

Una frase como «los expertos están de acuerdo» es más débil que una afirmación que identifica quién fue evaluado, qué pregunta se consideró, qué base de evidencia se revisó, cómo se trataron los desacuerdos y cuándo se hizo la evaluación. Los procesos formales de estudios de consenso son valiosos en parte porque hacen inspeccionable al menos una parte de ese recorrido. [8]

El mismo principio se aplica a la revisión. COPE insiste en conflictos declarados, experiencia relevante, confidencialidad y razonamiento constructivo porque una revisión gana legitimidad por cómo se realizó, no solo por la etiqueta «revisado por pares». [1]

Afirmación C10. Las afirmaciones sobre revisión o consenso son más auditables cuando el proceso, los participantes, el alcance y la base de evidencia son explícitos. [1][8]

En qué punto estamos

Las afirmaciones sobre revisión o consenso son más auditables cuando identifican participantes, alcance, procedimiento y base de evidencia. Eso no garantiza un buen resultado, pero hace inspeccionable el camino hacia el juicio. Cuando la pista de auditoría es visible aparece un límite más profundo: ¿pueden el acuerdo o el desacuerdo sustituir a la evidencia de primer orden que se está evaluando?

11. Lo que el desacuerdo no puede hacer

El desacuerdo por sí solo no puede convertir dos afirmaciones en igualmente probables. No puede sustituir evidencia ausente. Tampoco el consenso puede transformar mediciones débiles en mediciones fuertes. Tanto el acuerdo como el desacuerdo son información de segundo orden: dicen algo sobre cómo una comunidad ha procesado la evidencia, pero no sustituyen la evidencia de primer orden.

Afirmación C11 - Síntesis declarada. El acuerdo y el desacuerdo expertos son evidencia sobre el estado de la evaluación, no sustitutos de las observaciones, mediciones y argumentos que se evalúan.

En qué punto estamos

El acuerdo y el desacuerdo expertos son información de segundo orden sobre la evaluación, no sustitutos de observaciones, mediciones y argumentos. Se trata de una síntesis declarada y no de un teorema derivado de una sola fuente. La siguiente pregunta es práctica: si las señales sociales informan pero son limitadas, ¿cómo combinarlas con la evidencia de primer orden ante una afirmación controvertida?

12. Una regla práctica de lectura

Cuando encuentres una afirmación controvertida, haz cuatro preguntas:

1. **¿Cuál es la evidencia de primer orden?** ¿Qué se midió, observó o derivó?
2. **¿Cuál fue el proceso de revisión?** ¿Quién examinó la afirmación, con qué experiencia y qué posibilidad tuvo de desafiarla?
3. **¿Qué es exactamente lo acordado o discutido?** ¿Datos, método, modelo, interpretación causal, incertidumbre o juicio de valor?
4. **¿Qué cambiaría el balance?** ¿Una replicación, una nueva medición, un experimento discriminante, un análisis corregido o una nueva restricción teórica?

Ese procedimiento no garantiza la respuesta correcta. Evita que las palabras *revisado por pares*, *consenso* y *controversia* hagan más trabajo epistémico del que pueden sostener.

Afirmación C12 - Síntesis declarada. La fuerza de una afirmación de revisión o consenso debe evaluarse junto con la evidencia subyacente, el proceso y las posibles pruebas discriminantes.

En qué punto estamos

Un lector disciplinado pregunta qué se midió, cómo se revisó la afirmación, dónde reside el desacuerdo y qué evidencia podría discriminar entre alternativas. La regla no garantiza acertar; impide que las etiquetas institucionales hagan más trabajo epistémico del que soportan. La última pregunta es de alcance: ¿qué artículos KNO vecinos poseen los problemas de construcción de evidencia y reproducibilidad que aquí se dejan fuera?

13. Frontera y casos heredados

KNO-020 se ocupa de la **capa de evaluación social**: cómo la revisión, el consenso y el desacuerdo organizan la crítica una vez producida la evidencia. No sustituye a KNO-016, que pregunta cómo se construye la evidencia, ni a KNO-021, que tratará la apertura, la infraestructura de reproducibilidad y las prácticas que hacen posible comprobar.

Cinco artículos KNO anteriores delegaron aquí casos concretos de desacuerdo. Su disposición es deliberadamente limitada. Las disputas de KNO-010 sobre especie y rango permanecen como ejemplos de desacuerdo clasificatorio; las definiciones permisiva y restrictiva de medida de KNO-011, como desacuerdo definicional; el regreso del experimentador de KNO-012, como caso sobre controles e intervención; realismo frente a adecuación empírica de KNO-013, como disputa de interpretación de modelos; y la declaración de KNO-016 de que revisión por pares y consenso pertenecen aquí queda satisfecha por el tratamiento general del escrutinio social. Ninguno de esos casos se vuelve a litigar. Permanecen como ejemplos en sus artículos propietarios y KNO-020 aporta el marco general para interpretarlos.

En qué punto estamos

KNO-020 cierra los casos delegados de evaluación social sin duplicar su contenido científico: los artículos anteriores conservan sus disputas concretas y este artículo aporta el marco general para interpretar revisión, consenso y desacuerdo. KNO-016 posee la construcción de evidencia y KNO-021 la infraestructura de ciencia abierta y reproducibilidad. Fijadas esas fronteras, ya puede resumirse qué sabemos, qué ignoramos y qué sigue siendo revisable.

Lo que sabemos y cómo lo sabemos

- La revisión por pares estructurada puede revelar problemas antes de la publicación, pero los juicios de los revisores varían y no son certificados de verdad. [1][2][3]
- Las decisiones editoriales pueden incorporar recomendaciones de revisores en conflicto en vez de limitarse a contar votos. [4]
- La reproducibilidad y la replicabilidad amplían el escrutinio más allá de la publicación y pueden reforzar, matizar o cuestionar afirmaciones. [5]
- Los procesos formales de consenso científico pueden combinar revisión de evidencia, deliberación experta y revisión independiente. [8]
- Las correcciones pueden preservar un registro auditable de errores encontrados después de publicar. [7]

Lo que todavía no sabemos

No existe una conversión numérica universal entre «grado de acuerdo experto» y probabilidad de que una afirmación sea verdadera. Los campos también emplean modelos de revisión, estándares de evidencia y mecanismos de formación de consenso distintos. La fiabilidad de esos procesos depende de la pregunta, la evidencia, los incentivos y la independencia de las personas y métodos implicados.

Qué podría hacernos cambiar de opinión

La evidencia de que un modelo de revisión falla sistemáticamente al detectar errores importantes, que un proceso de consenso suprime contraevidencia relevante o que un modelo minoritario produce repetidamente mejores predicciones debería reducir la confianza en la evaluación predominante. A la inversa, pruebas independientes repetidas que eliminen alternativas plausibles deberían aumentarla.

Si solo recuerdas una cosa

Revisión, consenso y desacuerdo organizan la crítica; su valor epistémico depende de la evidencia, los procedimientos y las pruebas discriminantes que hacen visibles.

Artículos relacionados

- KNO-012 · Experiments and Variable Control
- KNO-013 · Models, Maps and Reality
- KNO-014 · Causality and Correlation
- KNO-015 · Uncertainty and Probability
- KNO-016 · How Evidence Is Built: Putting the Tools Together
- KNO-021 · Open Science and Reproducibility

Referencias

1. COPE Council. *COPE Ethical Guidelines for Peer Reviewers*. Version 2, September 2017. DOI: 10.24318/cope.2019.1.9.
2. Bornmann, Lutz. "Scientific peer review." *Annual Review of Information Science and Technology* 45(1): 197-245 (2011). DOI: 10.1002/aris.2011.1440450112.
3. Benda, Wim G.G.; Engels, Tim C.E. "The predictive validity of peer review: A selective review of the judgmental forecasting qualities of peers, and implications for innovation in science." *International Journal of Forecasting* 27(1): 166-182 (2011). DOI: 10.1016/j.ijforecast.2010.03.003.
4. *Nature Genetics*. "Peer Review." Políticas editoriales, página web vigente. Sin DOI; política viva del editor.
5. National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press, 2019. DOI: 10.17226/25303.
6. Ioannidis, John P. A. "Why Most Published Research Findings Are False." *PLOS Medicine* 2(8): e124 (2005). DOI: 10.1371/journal.pmed.0020124.

7. Ioannidis, John P. A. "Correction: Why Most Published Research Findings Are False." *PLOS Medicine* 19(8): e1004085 (2022). DOI: 10.1371/journal.pmed.1004085.
 8. National Academies of Sciences, Engineering, and Medicine. "Consensus Study Reports", descripción institucional de National Academies Press. Página institucional viva; sin DOI para la descripción.
-

KNO-021 · Ciencia abierta y reproducibilidad

KNO · Formas de conocer

Juan Seguro

v0.3 · ES

Todos los derechos reservados

Resumen del artículo

Las afirmaciones científicas resultan más fáciles de inspeccionar cuando son accesibles los materiales necesarios para comprenderlas, repetirlas o discutirlos. La ciencia abierta reúne prácticas que intentan hacer posible esa inspección: compartir datos y código cuando corresponde, describir los métodos con precisión, distinguir análisis exploratorios de confirmatorios, conservar versiones y hacer visibles las correcciones. Reproducibilidad y replicabilidad están relacionadas, pero no son idénticas. Un resultado puede ser reproducible computacionalmente y seguir siendo erróneo, mientras que un estudio independiente puede no replicar por razones que no se reducen a mala conducta o incompetencia. Este artículo pregunta qué partes de un resultado científico deben seguir siendo inspeccionables para que investigadores posteriores reconstruyan el análisis, prueben la afirmación con nuevos datos y localicen dónde entra el desacuerdo. Trata la apertura como infraestructura para la crítica, no como certificado de verdad.

En 60 segundos

La ciencia abierta no es un interruptor llamado «compartirlo todo». Es un conjunto de controles que vuelve más inspeccionable el trabajo científico. La reproducibilidad suele preguntar si los mismos datos y procedimientos computacionales producen el mismo resultado; la replicabilidad pregunta si nuevos datos obtenidos para la misma cuestión generan resultados compatibles dentro de la incertidumbre pertinente.[1,2] El preregistro puede separar predicciones de análisis escogidos tras ver los datos, pero no convierte un diseño débil en uno fuerte.[3] Los informes registrados trasladan parte de la evaluación a antes de conocer los resultados, mientras que los datos y el código abiertos permiten inspeccionar transformaciones, exclusiones e implementación.[4] Ninguna de estas prácticas garantiza la verdad. Su valor es procedimental: exponen más del camino entre observación y afirmación, conservan alternativas que de otro modo desaparecerían y abaratan la corrección. La apertura también tiene límites: privacidad, seguridad, consentimiento, propiedad intelectual y carga de curación pueden justificar acceso controlado. La meta es ciencia auditable, no divulgación sin fricción.

Por qué importa esta pregunta

La ciencia moderna depende a menudo de conjuntos de datos, scripts, instrumentos y decisiones que no aparecen en el artículo final. Cuando solo sobrevive la narración, investigadores posteriores pueden no saber si un desacuerdo nace de los datos, el preprocesamiento, las decisiones estadísticas, el software, la interpretación o una diferencia real en la naturaleza. Las prácticas de ciencia abierta intentan conservar esos puntos de decisión. Importan porque la crítica es uno de los mecanismos de corrección de errores de la ciencia, pero solo puede operar sobre lo inspeccionable. Las mismas prácticas revelan además una tensión: la apertura puede aumentar la auditabilidad y, al mismo tiempo, crear nuevas obligaciones de privacidad, gobernanza y conservación a largo plazo.

1. Reproducibilidad no es replicación

La pregunta Si dos investigadores obtienen el mismo resultado numérico, ¿qué se ha reproducido exactamente: el cálculo, la observación o la afirmación? La respuesta depende de qué entradas, procedimientos y condiciones permanecen fijos y cuáles pueden cambiar durante la nueva prueba.

Las National Academies distinguen reproducibilidad de replicabilidad para ganar claridad práctica. La reproducibilidad consiste en obtener resultados consistentes con los mismos datos de entrada, pasos computacionales, métodos y condiciones de análisis; la replicabilidad se refiere a nuevos estudios sobre la misma pregunta con datos obtenidos de nuevo.[1] Goodman, Fanelli e Ioannidis muestran igualmente que el vocabulario se ha usado de forma inconsistente entre campos y advierten contra tratar «reproducible» como sinónimo de «verdadero».[2]

La distinción separa dos modos de fallo. Si los mismos archivos y código no regeneran una tabla publicada, el problema puede ser una dependencia ausente, preprocesamiento no documentado, deriva de versiones o error de implementación. Si un estudio nuevo discrepa, la explicación puede estar en muestreo, medición, contexto, heterogeneidad del efecto o una afirmación original falsa. El diagnóstico empieza preguntando qué tipo de igualdad se esperaba.

Afirmación C01. La reproducibilidad y la replicabilidad ponen a prueba formas distintas de consistencia. [1,2]

Afirmación C10. La reproducibilidad no debe usarse como sinónimo de verdad. [2]

En qué punto estamos

Reproducibilidad y replicación prueban eslabones distintos de la cadena evidencial. Repetir un cálculo antiguo comprueba si el análisis comunicado puede reconstruirse; obtener nuevos datos prueba si la afirmación sobrevive a un nuevo encuentro con el mundo. Ningún resultado es certificado de verdad. La pregunta siguiente es qué debe conservarse para que la reconstrucción sea posible.

2. Un artículo no es todo el análisis

Un artículo comprime un proceso mucho mayor. Las observaciones brutas se seleccionan, calibran, limpian, transforman y resumen antes de convertirse en figura o coeficiente. Las bibliotecas de software aportan valores por defecto; las reglas de exclusión eliminan casos; los archivos intermedios conservan estado. Cuando esas capas desaparecen, la publicación puede seguir siendo legible mientras el análisis deja de ser reconstruible.

Un trabajo reproducible necesita por tanto más que prosa. Según el campo, puede requerir datos o un sustituto de acceso controlado, código, información del entorno, parámetros, registros de calibración y un mapa que conecte salidas con entradas. No se trata de conservar cada pulsación, sino de preservar suficiente grafo de dependencias para que otro investigador competente identifique cómo se produjo el resultado.[1,4]

Afirmación C02. La reproducibilidad puede fallar cuando los datos, el código, las dependencias o las condiciones del análisis no pueden reconstruirse. [1]

En qué punto estamos

El objeto científico sometido a revisión suele ser mayor que el artículo que lo describe. Datos, código, parámetros y procedencia pueden ser responsables del resultado aunque no ocupen ninguna página impresa. Conservarlos vuelve inspeccionables decisiones ocultas. Eso conduce al siguiente problema: algunas decisiones deben registrarse antes de conocer el resultado, no reconstruirse después.

3. El prerregistro separa predicción y retrospección

Los datos pueden sugerir hipótesis después de ser observados. Eso es científicamente útil, pero no equivale a probar una predicción que existía antes de la observación. Nosek, Ebersole, DeHaven y Mellor describen el prerregistro como una forma de hacer visible esa diferencia temporal registrando hipótesis y planes de análisis antes de inspeccionar los resultados.[3]

El prerregistro no elimina la exploración ni vuelve correcto un análisis. Los planes pueden ser imprecisos, las mediciones pueden fallar y pueden ser necesarias desviaciones justificadas. Su valor es que un lector posterior puede distinguir qué decisiones fueron prospectivas y cuáles respondieron a los datos. Eso cambia cómo se interpreta la fuerza evidencial sin prohibir el descubrimiento.

Afirmación C03. El prerregistro puede hacer visible el momento en que se formularon las hipótesis y se eligieron los análisis. [3]

Afirmación C04 - Síntesis declarada. El prerregistro no establece por sí solo la calidad del diseño ni la verdad. **Síntesis declarada:** [3] respalda el prerregistro como compromiso procedimental fijado antes de conocer los resultados; la cláusula adicional de que no establece por sí mismo la calidad del diseño ni la verdad es la síntesis.

En qué punto estamos

El prerregistro es una distinción con marca temporal, no un sello de calidad. Conserva cuándo entró una hipótesis o decisión analítica en el flujo de trabajo, permitiendo separar predicción de adaptación dependiente de los datos. Puede reducir la ambigüedad del estatus evidencial sin deslegitimar la exploración. El siguiente paso lleva esa lógica a la publicación: ¿puede empezar la evaluación antes de conocer el resultado?

4. Los informes registrados cambian el orden de la evaluación

Un artículo convencional suele revisarse cuando ya existen los resultados principales. Los informes registrados adelantan una parte importante de la revisión: la pregunta, el diseño y el análisis previsto pueden evaluarse antes de conocer el desenlace. Munafò y colegas incluyen este mecanismo entre reformas destinadas a reducir oportunidades de comunicación selectiva y narración retrospectiva.[4]

Cambiar el orden no garantiza una pregunta útil ni una ejecución impecable. Cambia incentivos y el registro evidencial. Un resultado nulo procedente de un diseño aceptado en principio tiene menos riesgo de desaparecer por ser poco llamativo, mientras que las desviaciones del plan pueden documentarse en lugar de ocultarse. La publicación queda más ligada a la calidad de pregunta y método que a la dirección del resultado.

Afirmación C05. Las reformas del tipo informe registrado desplazan parte de la evaluación de preguntas y métodos a una fase anterior del proceso de investigación. [4]

En qué punto estamos

Los informes registrados muestran que la apertura también trata de cronología: conservar qué se decidió antes de que los resultados pudieran influir en esas decisiones. Reducen una vía por la que el registro publicado puede quedar filtrado selectivamente, pero no eliminan todos los filtros. La pregunta siguiente es qué ocurre después de publicar, cuando otro investigador necesita acceder al material que hizo posible el resultado.

5. Datos y código abiertos vuelven inspeccionables las transformaciones

Compartir datos permite comprobar codificación, exclusiones, transformaciones y análisis alternativos. Compartir código puede mostrar cómo se calculó realmente un número publicado, no solo cómo se describió el método en prosa. Juntos reducen la distancia entre la afirmación y las operaciones que la produjeron.[1,4]

Pero «abierto» no basta. Un archivo sin definiciones de variables, procedencia o versión puede ser descargable y científicamente inútil. El código que depende de versiones de software no documentadas puede fallar años después. La apertura necesita curación: identificadores, metadatos, licencias, información de dependencias y estrategias de conservación. KNO-019 explicó por qué los registros estructurados necesitan contexto; la ciencia abierta añade que ese contexto debe seguir accesible a la crítica.

Afirmación C06. Los datos y el código abiertos pueden exponer a escrutinio transformaciones y decisiones de implementación. [1,4]

Afirmación C07 - Síntesis declarada. El acceso abierto sin metadatos, procedencia y contexto de versión puede ser científicamente insuficiente. **Síntesis declarada:** [1] respalda la importancia de metadatos, procedencia y contexto de versión para la reutilización; el juicio de que el acceso por sí solo puede ser científicamente insuficiente es la síntesis.

En qué punto estamos

Los materiales abiertos extienden el objeto inspeccionable desde el artículo hasta el flujo de trabajo, pero el acceso sin contexto es una infraestructura débil. Datos y código necesitan metadatos, versiones y procedencia para seguir siendo interpretables. La apertura se convierte así en un problema de custodia además de divulgación. También aparece un límite: algunos materiales científicamente importantes no pueden hacerse públicos responsablemente sin restricciones.

6. La apertura tiene límites legítimos

Los datos de participantes humanos pueden contener información que no fue consentida para publicación. Localizaciones ecológicas pueden exponer especies amenazadas. Métodos sensibles para la seguridad y obligaciones contractuales pueden restringir la difusión. Por eso ciencia abierta no puede significar publicación pública universal de todo objeto subyacente.

El principio más fuerte es inspeccionabilidad bajo gobernanza apropiada. Desidentificación, entornos seguros, acuerdos de uso, datos sintéticos, repositorios controlados y comités independientes pueden conservar parte de la auditabilidad sin fingir que el riesgo es cero. El problema de diseño consiste en hacer las restricciones explícitas y proporcionales para que «no puede ser público» no se convierta silenciosamente en «no puede comprobarse».

Afirmación C08. Las obligaciones legítimas de privacidad y seguridad pueden justificar un acceso controlado en vez de irrestricto. [1]

En qué punto estamos

La ciencia abierta es compatible con acceso controlado cuando la difusión irrestricta vulneraría obligaciones legítimas. El requisito epistémico no es publicidad máxima, sino una vía documentada para que el escrutinio cualificado siga siendo posible cuando sea viable. Eso convierte la gobernanza en parte de la infraestructura evidencial. La cuestión siguiente es la corrección: cuando se detecta un error, ¿queda registrado que el objeto científico cambió?

7. Versiones y correcciones forman parte de la evidencia

Un conjunto de datos corregido, un script revisado o una conclusión retractada no son el mismo objeto científico que la versión anterior al cambio. Identificadores persistentes, repositorios versionados y avisos de corrección permiten recuperar esa distinción. Sin ellos, usuarios posteriores pueden mezclar sin saberlo afirmaciones de una versión con datos o código de otra.

Una buena corrección hace más que sustituir un archivo. Conserva suficiente historia para mostrar qué cambió, por qué cambió y qué conclusiones quedan afectadas. Es una forma de memoria institucional. Permite que la ciencia cambie de opinión sin borrar el camino del cambio. KNO-024 volverá al problema de la conservación desde la dirección opuesta: qué desaparece cuando esos registros no sobreviven.

Afirmación C09 - Síntesis declarada. El versionado y los registros de correcciones distinguen objetos científicos modificados de sus versiones anteriores. **Síntesis declarada:** [1] respalda hacer explícitos el contexto de versión y las correcciones; tratar esos registros como lo que distingue un objeto científico modificado de su manifestación anterior es la síntesis.

En qué punto estamos

El versionado hace visible la corrección como cambio del objeto evidencial y no como sustitución silenciosa. Importa porque investigadores posteriores deben saber a qué datos, código e interpretación se refiere una afirmación. Un historial transparente de correcciones es evidencia sobre la historia de la evidencia. La pregunta final es cómo juzgar estas prácticas sin convertir la apertura en una insignia de corrección.

8. La ciencia abierta es infraestructura de corrección de errores

Las prácticas abiertas no garantizan que un conjunto de datos carezca de sesgo, que un modelo esté bien especificado o que una conclusión sea verdadera. Cambian qué errores pueden encontrarse y quién puede encontrarlos. Munafò y colegas presentan la práctica reproducible como un sistema de

reformas que interactúan, no como una cura única.[4] Las National Academies tratan igualmente reproducibilidad y replicabilidad como propiedades que deben investigarse, no como veredictos simples sobre el valor científico.[1]

El argumento más fuerte a favor de la apertura es por tanto procedimental. Una afirmación acompañada de materiales inspeccionables expone más puntos posibles de crítica. Un resultado que no replica puede diagnosticarse con mayor precisión. Una corrección puede propagarse a los objetos que dependen de ella. La ganancia no es certeza, sino un camino más corto desde el desacuerdo al diagnóstico y del diagnóstico a la reparación.

Afirmación C11 - Síntesis declarada. Las prácticas de ciencia abierta pueden acortar el camino desde el desacuerdo hasta el diagnóstico al exponer una parte mayor del flujo evidencial. **Síntesis declarada:** [1,4] respaldan abrir registros y flujos de trabajo a inspección y reutilización; la afirmación de que eso puede acortar el camino del desacuerdo al diagnóstico es la síntesis.

Afirmación C12 - Síntesis declarada. La ciencia abierta se entiende mejor como infraestructura de corrección de errores que como certificado de verdad. **Síntesis declarada:** [1,2,4] respaldan transparencia, prácticas de reproducibilidad y mecanismos de corrección; describir la ciencia abierta en conjunto como infraestructura de corrección de errores y no como certificado de verdad es la síntesis.

En qué punto estamos

La ciencia abierta se entiende mejor como infraestructura para la crítica organizada. Conserva más del recorrido entre observación y afirmación, distingue algunas decisiones por cuándo se tomaron y facilita rastrear correcciones. Estas propiedades pueden mejorar la fiabilidad sin demostrar verdad. La frontera queda clara: KNO-020 posee revisión y consenso; KNO-021 posee inspeccionabilidad, reproducibilidad y prácticas de corrección.

Lo que sabemos y cómo lo sabemos

- Reproducibilidad y replicabilidad prueban tipos distintos de consistencia y no deben tratarse como sinónimos.[1,2]
- El prerregistro puede hacer visible el momento en que se formularon hipótesis y decisiones analíticas.[3]
- Datos, código y procedencia abiertos pueden volver inspeccionable una parte mayor del análisis, pero el acceso sin contexto es insuficiente.[1,4]
- Versionado y correcciones visibles conservan a qué objeto científico se refiere una afirmación.

Lo que todavía no sabemos

- Cuánto mejora cada intervención de ciencia abierta la fiabilidad en disciplinas y diseños distintos.
- Qué combinaciones de apertura y acceso controlado preservan mejor auditabilidad y privacidad legítima.
- Cuánto tiempo siguen siendo ejecutables los objetos de investigación actuales cuando cambian software y hardware.
- Cómo alteran los incentivos el uso real, y no solo la adopción nominal, de prácticas abiertas.

Qué podría hacernos cambiar de opinión

- Evaluaciones amplias entre disciplinas que muestren que una práctica promovida no añade auditabilidad ni fiabilidad en uso real.
- Evidencia de que un control destinado a reducir comunicación selectiva crea distorsiones mayores de las que evita.
- Nuevos métodos de preservación de privacidad que vuelvan rutinarias y seguras formas hoy restringidas de escrutinio.
- Estudios de conservación a largo plazo que muestren que repositorios y versionado actuales no mantienen interpretables los objetos de investigación.

Si solo recuerdas una cosa

La ciencia abierta no certifica la verdad; conserva más del camino evidencial para que otros puedan inspeccionarlo, reproducirlo, discutirlo, corregirlo y reutilizarlo.

Artículos relacionados

- KNO-019 · Datos, tablas y registros
- KNO-020 · Revisión, consenso y desacuerdo
- KNO-024 · Conocimiento perdido y sesgo de conservación

Referencias

[1] National Academies of Sciences, Engineering, and Medicine. Reproducibility and Replicability in Science. National Academies Press, Washington, DC (2019). DOI 10.17226/25303. <https://doi.org/10.17226/25303> [2] Steven N. Goodman; Daniele Fanelli; John P. A. Ioannidis. What does research reproducibility mean?. Science Translational Medicine 8(341), 341ps12 (2016). DOI 10.1126/scitranslmed.aaf5027. <https://doi.org/10.1126/scitranslmed.aaf5027> [3] Brian A. Nosek; Charles R. Ebersole; Alexander C. DeHaven; David T. Mellor. The preregistration revolution. Proceedings of the National Academy of Sciences 115(11), 2600–2606 (2018). DOI 10.1073/pnas.1708274114. <https://doi.org/10.1073/pnas.1708274114> [4] Marcus R. Munafò; Brian A. Nosek; Dorothy V. M. Bishop; Katherine S. Button; Christopher D. Chambers; Nathalie Percie du Sert; Uri Simonsohn; Eric-Jan Wagenmakers; Jennifer J. Ware; John P. A. Ioannidis. A manifesto for reproducible science. Nature Human Behaviour 1, 0021 (2017). DOI 10.1038/s41562-016-0021. <https://doi.org/10.1038/s41562-016-0021>

KNO-022 · La computación como memoria externa

KNO · Formas de conocer

Juan Segurola

v0.3 · ES

Todos los derechos reservados

Resumen del artículo

Un ordenador hace más que calcular deprisa. Puede almacenar estados intermedios, aplicar procedimientos explícitos repetidamente, conservar estructuras simbólicas y hacer operaciones inspeccionables fuera de los límites de la memoria humana sin ayuda. Eso convierte la computación en infraestructura cognitiva externa. Sin embargo, externalizar memoria y procedimiento cambia qué recuerdan las personas, qué observan y dónde pueden ocultarse los errores. Un archivo guardado puede conservar un resultado y perder el entorno que lo produjo; una tubería automatizada puede aumentar consistencia y repetir el mismo error miles de veces. Este artículo pregunta qué aporta la computación cuando se usa como memoria externa para ciencia y razonamiento: persistencia, repetibilidad, escala, representaciones compartibles y procedimientos ejecutables. También pregunta qué no puede aportar por sí sola: variables con significado, supuestos justificados, entradas fiables o una interpretación de por qué importa el resultado.

En 60 segundos

Los humanos trasladan rutinariamente trabajo cognitivo al entorno: escribimos listas, dibujamos diagramas, ponemos recordatorios y ordenamos objetos para que el mundo cargue con parte de la tarea. La investigación sobre descarga cognitiva trata esto como estrategia general, no como anomalía.[1] Los ordenadores amplían la misma lógica. Pueden conservar estados simbólicos exactos, repetir procedimientos y transformar representaciones a escalas que la memoria sin ayuda no puede manejar. Las representaciones externas cambian además la estructura del pensamiento, no solo su carga de almacenamiento.[2] Motores de búsqueda e información en red pueden desplazar la memoria hacia saber dónde encontrar la información.[3] En ciencia, el código ejecutable puede conservar un procedimiento con más precisión que la prosa, sobre todo si documentación y cómputo permanecen unidos.[4] Pero la computación no es memoria neutral. Formatos, versiones, interfaces y automatización modelan qué sobrevive y qué resulta fácil. Un ordenador puede conservar perfectamente un error. Su valor epistémico surge al acoplar memoria externa con procedencia, interpretación y controles.

Por qué importa esta pregunta

Muchos resultados científicos dependen hoy de cadenas de cómputo demasiado largas para reproducirse mentalmente y demasiado detalladas para describirse por completo en prosa. La cuestión práctica ya no es si los humanos usan memoria externa, sino cuán fiable es esa memoria. Scripts, cuadernos, bases de datos e historiales de versiones pueden conservar pasos de razonamiento entre tiempos y personas. También pueden separar a los usuarios de los supuestos incorporados en esos pasos. Entender la computación como memoria externa ayuda a distinguir tres cosas que suelen confundirse: almacenar información, ejecutar un procedimiento y comprender qué significa ese procedimiento.

1. La memoria puede trasladarse al entorno

La pregunta ¿Qué cambia cuando la información necesaria para una tarea se almacena fuera de la persona que la realiza? La memoria externa puede reducir la carga interna, pero también reorganizar el razonamiento, crear nuevas dependencias y cambiar lo que el usuario debe saber para recuperar, ejecutar e interpretar la información almacenada.

La descarga cognitiva usa acciones físicas o almacenamiento externo para reducir demanda cognitiva interna. Risko y Gilbert revisan evidencia de que las personas reestructuran tareas escribiendo, ordenando o almacenando información cuando depender solo de la memoria sería costoso.[1] La ganancia es persistencia y posibilidad de compartir, no inmunidad al error: un registro externo puede perderse, etiquetarse mal o quedar sin contexto.

Afirmación C01. Las personas usan acciones y representaciones externas para reducir la demanda cognitiva interna. [1,2]

En qué punto estamos

La memoria externa cambia quién puede inspeccionar información, cuánto puede persistir y qué errores se vuelven posibles. Escribir y almacenar reduce algunas cargas cognitivas y crea dependencia de etiquetas, ubicación y contexto. La computación amplía este intercambio al volver ejecutables los estados externos. La pregunta siguiente es cómo las representaciones remodelan el razonamiento en vez de limitarse a guardar hechos.

2. Las representaciones externas cambian el coste de pensar

Kirsh sostiene que las representaciones externas hacen más que ahorrar memoria: pueden cambiar el coste de la inferencia, crear referentes persistentes y ofrecer objetos compartibles de pensamiento.[2] Un diagrama, tabla o disposición simbólica puede hacer simultáneamente disponibles relaciones difíciles de mantener internamente. La computación añade la capacidad de transformar una representación externa en otra mediante reglas explícitas.

Afirmación C02. Las representaciones externas pueden alterar la estructura de costes del razonamiento, no solo la carga de memoria. [2]

En qué punto estamos

Las representaciones externas reorganizan operaciones, no solo almacenamiento. Una tabla, gráfica o estructura simbólica cambia qué puede compararse a la vez y qué relaciones resultan fáciles de inspeccionar. La computación añade transformación regida por reglas entre representaciones. Eso conduce a una memoria externa más fuerte: un estado almacenado puede emparejarse con un procedimiento ejecutable que lo cambia de forma predecible.

3. Un programa conserva un procedimiento, no solo un resultado

Una respuesta guardada registra un punto final; un programa puede conservar una secuencia de operaciones capaz de regenerarlo desde entradas especificadas. La propuesta de programación literaria de Knuth hizo explícito el vínculo entre procedimiento ejecutable y explicación humana.[4] El código puede externalizar memoria procedimental con más detalle operativo que la prosa, aunque sigue haciendo falta documentación que conecte operaciones de máquina con significado científico.

Afirmación C03. El código puede conservar un procedimiento ejecutable además de un resultado almacenado. [4]

Afirmación C10. La programación literaria vincula el procedimiento ejecutable con una explicación legible por humanos. [4]

En qué punto estamos

El código puede externalizar memoria procedimental: no solo qué respuesta se obtuvo, sino cómo un sistema puede volver a obtenerla desde entradas. Eso aumenta precisión y repetibilidad en flujos complejos si el código sigue siendo interpretable. La ejecución por sí sola no revela por qué una variable representa lo que afirma representar. La siguiente pregunta es qué ocurre cuando el procedimiento recordado depende de un entorno cambiante.

4. Archivos exactos pueden dejar de ser ejecutables

Un script puede permanecer idéntico byte a byte mientras las bibliotecas cambian valores por defecto, los sistemas operativos eliminan interfaces, desaparecen servicios y envejecen los formatos. El texto de un programa no especifica por completo su comportamiento futuro si no se conocen dependencias y entorno. Contenedores y manifiestos reducen ambigüedad, pero la memoria computacional duradera sigue requiriendo migración y pruebas, no solo copias.

Afirmación C04 - Síntesis declarada. La preservación ejecutable depende del contexto de software y entorno además de la identidad del archivo. **Síntesis declarada:** [4] respalda conservar información del entorno computacional; afirmar que la preservación ejecutable depende por ello de ese contexto además de la identidad del archivo es la síntesis.

En qué punto estamos

La persistencia digital tiene capas. La identidad de bytes protege frente a un tipo de cambio, mientras la ejecutabilidad depende del contexto de software y hardware, y el significado científico de metadatos e interpretación. Un script conservado puede estar intacto y ser inutilizable. Esto revela el siguiente límite de la automatización: un procedimiento fiable puede repetir un supuesto erróneo con extraordinaria consistencia.

5. La automatización multiplica consistencia y error

La automatización aplica repetidamente el mismo procedimiento explícito sin fatiga y permite procesar volúmenes que una persona no podría manejar manualmente. Esa consistencia depende de que el procedimiento sea apropiado. Un error de conversión de unidades dentro de un bucle no se diluye al repetirlo; se replica. La escala computacional aumenta por tanto el valor de pruebas, datos de validación, invariantes y controles independientes, porque el error sistemático puede amplificarse con regularidad perfecta.

Afirmación C05 - Síntesis declarada. La automatización puede reducir la variabilidad de ejecución y a la vez amplificar un error procedimental sistemático. **Síntesis declarada:** no se asigna ninguna fuente como apoyo directo; tanto la reducción de variabilidad entre ejecuciones como la posibilidad de repetir de forma consistente el mismo error procedimental son aquí síntesis metodológica.

En qué punto estamos

La automatización cambia la distribución del error. Puede reducir inconsistencia producida por ejecución manual repetida y magnificar cualquier fallo incrustado en el procedimiento. Cuanto más se reutiliza una tubería, más importantes se vuelven sus supuestos. Por eso la memoria computacional es inseparable de la verificación: los procedimientos almacenados necesitan pruebas que puedan fallar. La pregunta siguiente es qué recuerdan los usuarios cuando las máquinas recuerdan por ellos.

6. Saber dónde puede sustituir a recordar qué

Sparrow, Liu y Wegner describieron experimentos en los que la expectativa de acceso posterior por ordenador afectó a lo que los participantes recordaban, incluida la ubicación de la información.[3] El resultado no muestra que internet simplemente destruya memoria. Ilustra un cambio de asignación: cuando se espera recuperación externa, la memoria interna puede priorizar rutas de acceso frente al contenido, creando dependencia de la estabilidad del sistema externo.

Afirmación C06. La expectativa de disponer después de acceso informático puede cambiar qué recuerdan las personas sobre la información y su ubicación. [3]

Afirmación C09 - Síntesis declarada. La memoria computacional externa crea dependencia de identificadores, interfaces e instituciones. **Síntesis declarada:** [1,3] respaldan la cognición distribuida y los recursos computacionales persistentes; describir su uso como dependencia de identificadores, interfaces e instituciones es la síntesis.

En qué punto estamos

El acceso externo puede cambiar para qué se usa la memoria interna. Las personas pueden conservar rutas hacia la información en vez de la información misma cuando esperan recuperación fiable. Es eficiente mientras persiste el sistema externo, pero crea dependencia de identificadores, interfaces e instituciones. La pregunta siguiente es cómo se coordinan esas dependencias cuando la computación se distribuye entre personas, artefactos y procedimientos.

7. La computación puede distribuirse entre personas y artefactos

El estudio de Hutchins sobre navegación trata el trabajo cognitivo como distribuido entre personas, instrumentos, cartas, procedimientos y comunicación, no localizado por completo en un navegante.[5] Los flujos científicos modernos tienen arquitectura similar: sensores producen registros, scripts los transforman, bases de datos conservan estado y especialistas validan etapas. La fiabilidad puede emerger de la división coordinada del trabajo, mientras responsabilidad y significado pueden diluirse en los traspasos.

Afirmación C07. Los sistemas cognitivos distribuidos pueden abarcar personas, artefactos y procedimientos. [5]

Afirmación C08. Ningún participante individual necesita representar cada transformación de un flujo cognitivo distribuido. [5]

En qué punto estamos

El conocimiento computacional puede distribuirse entre humanos y artefactos. Ningún individuo necesita recordar cada estado intermedio si el sistema los conserva y coordina transformaciones con fiabilidad. Esto amplía la capacidad colectiva y vuelve más importante la procedencia: responsabilidad y significado pueden perderse en los traspasos. KNO-023 examinará el lado humano de esta arquitectura; antes, KNO-022 debe fijar su frontera con la comprensión misma.

8. La memoria externa no aporta significado

Licklider imaginó una asociación en la que las máquinas realizan trabajo rutinizable mientras los humanos formulan objetivos, hipótesis, criterios y evaluaciones.[6] El lenguaje es antiguo, pero la advertencia metodológica sigue siendo útil. Un ordenador puede conservar y ejecutar operaciones formales sin decidir si las entradas responden a la pregunta correcta. El significado sigue entrando mediante decisiones de modelado, convenciones de medida e interpretación.

Afirmación C11 - Síntesis declarada. La computación puede conservar operaciones sin establecer el significado científico de sus entradas. **Síntesis declarada:** [4,6] respaldan conservar flujos de trabajo y operaciones dependientes de software; el contraste de que esas operaciones conservadas no establecen por sí solas el significado científico de las entradas es la síntesis.

Afirmación C12 - Síntesis declarada. La computación es epistémicamente más fuerte cuando los estados y procedimientos almacenados conservan procedencia e interpretación. **Síntesis declarada:** [2,4,5] respaldan procedencia, procedimientos reproducibles y contexto computacional conservado; considerar epistémicamente más fuertes los sistemas que mantienen esos elementos interpretables es la síntesis.

En qué punto estamos

La computación es memoria externa más transformación ejecutable, no una fuente autónoma de significado. Puede conservar estados, procedimientos y coordinación a través del tiempo, pero la importancia científica de esos objetos sigue dependiendo de representaciones y preguntas justificadas. La frontera queda fijada: KNO-022 posee descarga computacional y memoria ejecutable; KNO-023 pasa a cómo los grupos humanos distribuyen y combinan conocimiento.

Lo que sabemos y cómo lo sabemos

- Las personas usan representaciones y acciones externas para reducir demanda cognitiva interna.[1,2]
- Las representaciones externas pueden cambiar estructura y coste del razonamiento, no solo necesidades de almacenamiento.[2]
- La expectativa de acceso externo puede afectar qué se recuerda de la información y su ubicación.[3]
- El código ejecutable puede conservar procedimientos con más especificidad operativa que la prosa por sí sola.[4]

Lo que todavía no sabemos

- Cómo distintas formas de descarga cognitiva remodelan aprendizaje a largo plazo y no solo rendimiento inmediato.
- Qué dependencias computacionales deben conservarse para ejecutabilidad a escala de siglos.
- Cuánto cambia la automatización la capacidad de detectar errores conceptuales y no solo de implementación.
- Dónde debe situarse la responsabilidad en flujos humano-computador muy distribuidos.

Qué podría hacernos cambiar de opinión

- Evidencia de que las representaciones externas perjudican sistemáticamente el razonamiento incluso controlando demanda y costes de recuperación.
- Estudios a largo plazo que muestren que herramientas actuales de reproducibilidad no preservan flujos ejecutables pese a metadatos completos.
- Demostraciones de que los errores de automatización se detectan fiablemente sin validación independiente, reduciendo la necesidad de controles redundantes.
- Nuevos modelos cognitivos que expliquen la descarga sin papel funcional de memoria o representación externa.

Si solo recuerdas una cosa

La computación gana poder epistémico cuando conserva no solo resultados, sino también estados, procedimientos y procedencia necesarios para entender cómo se produjeron.

Artículos relacionados

- KNO-008 · Instrumentos que amplían los sentidos
- KNO-019 · Datos, tablas y registros
- KNO-021 · Ciencia abierta y reproducibilidad

Referencias

[1] Evan F. Risko; Sam J. Gilbert. Cognitive Offloading. Trends in Cognitive Sciences 20(9), 676–688 (2016). DOI 10.1016/j.tics.2016.07.002. <https://doi.org/10.1016/j.tics.2016.07.002> [2] David Kirsh. Thinking with external representations. AI & Society 25, 441–454 (2010). DOI 10.1007/s00146-010-0272-8. <https://doi.org/10.1007/s00146-010-0272-8> [3] Betsy Sparrow; Jenny Liu; Daniel M. Wegner. Google effects on memory: cognitive consequences of having information at our fingertips. Science 333(6043), 776–778 (2011). DOI 10.1126/science.1207745. <https://doi.org/10.1126/science.1207745> [4] D. E. Knuth. Literate Programming. The Computer Journal 27(2), 97–111 (1984). DOI 10.1093/comjnl/27.2.97. <https://doi.org/10.1093/comjnl/27.2.97> [5] Edwin Hutchins. Cognition in the Wild. The MIT Press (1995). DOI 10.7551/mitpress/1881.001.0001. <https://doi.org/10.7551/mitpress/1881.001.0001> [6] J. C. R. Licklider. Man-Computer Symbiosis. IRE Transactions on Human Factors in Electronics HFE-1, 4–11 (1960).

KNO-023 · Cómo los grupos saben más que los individuos

KNO · Formas de conocer

Juan Segurola

v0.3 · ES

Todos los derechos reservados

Resumen del artículo

Los grupos a veces resuelven problemas que superan a cualquiera de sus miembros, pero añadir personas no añade automáticamente conocimiento. El rendimiento colectivo depende de cómo se distribuye la información, de si los errores son independientes, de cómo se expresa el desacuerdo y de cómo se combinan las aportaciones. Perspectivas diversas pueden superar, bajo determinadas condiciones de modelo, a grupos seleccionados solo por capacidad individual, mientras que la influencia social puede reducir precisamente la diversidad que hace útil la agregación.[1,2] En laboratorio, algunos grupos también superan a su mejor miembro en tareas intelectivas concretas.[3] Este artículo pregunta cuándo un grupo se convierte en sistema epistémico y no solo en multitud: cuando los miembros poseen información complementaria, cuando la estructura permite que aflore, cuando pueden compararse razones y cuando la decisión final conserva en vez de borrar el desacuerdo informativo. Trata la inteligencia colectiva como organización condicional, no como propiedad que aparece solo porque haya muchas personas.

En 60 segundos

Un grupo puede ganar poder epistémico de varias formas. Distintos miembros pueden conocer hechos distintos; estimaciones independientes pueden cancelar parte de los errores individuales; la discusión puede exponer contradicciones; roles especializados pueden conservar conocimiento que ninguna persona posee completo. Experimentos muestran que algunos grupos superan a su mejor miembro en tareas estructuradas de resolución de problemas.[3] Los modelos de Hong y Page muestran cómo puede importar la diversidad de enfoques bajo supuestos especificados.[1] Pero la agregación puede fallar. Lorenz y colegas encontraron que la influencia social puede reducir la diversidad de estimaciones sin mejorar necesariamente la precisión colectiva.[2] Un grupo también puede amplificar sesgos compartidos, suprimir información minoritaria, seguir prestigio en lugar de evidencia o converger antes de explorar alternativas. La inteligencia colectiva exige por tanto arquitectura: mecanismos para obtener información independiente, asignar experiencia, registrar disenso, integrar aportaciones y revisar decisiones. La pregunta clave no es si los grupos son más sabios en general, sino qué estructuras preservan la información que los individuos poseen colectivamente.

Por qué importa esta pregunta

Ciencia, ingeniería, medicina e instituciones públicas dependen de grupos que deben combinar conocimiento distribuido entre personas. Ningún individuo lee todos los artículos, realiza todas las mediciones ni comprende todos los subsistemas. La fiabilidad del conocimiento colectivo depende por tanto del diseño organizativo tanto como de la experiencia individual. Un comité puede superar a cualquiera de sus miembros al combinar información complementaria o rendir peor al forzar acuerdo

prematureo. Entender la diferencia importa porque el consenso suele tratarse como un simple recuento de cabezas de acuerdo. KNO-023 pregunta qué estructura de procesamiento de información produjo ese acuerdo.

1. Más personas no significan automáticamente más conocimiento

La pregunta ¿Qué debe ocurrir entre poseer información distinta y producir un mejor juicio colectivo? Más personas pueden aportar evidencia, rutas de búsqueda y crítica, pero la interacción también puede borrar independencia, amplificar el estatus y colapsar la diversidad antes de que se pongan a prueba alternativas útiles.

Un grupo contiene más experiencias y observaciones totales que un miembro típico, pero esos recursos solo importan si entran en la decisión. Un sesgo compartido puede hacer que diez personas repitan el mismo error, mientras un reparto desigual de la palabra puede ocultar información única. La unidad útil no es el número de cabezas, sino la estructura de información: quién sabe qué, cuán independientes son las piezas y cómo se combinan.

Afirmación C01 - Síntesis declarada. El rendimiento de un grupo depende de la estructura de la información, no solo del número de integrantes. **Síntesis declarada:** [1-4] respaldan efectos de diversidad, influencia social, tamaño del grupo y proceso de interacción; combinarlos en la regla general de que importa más la estructura informativa que el simple número es la síntesis.

En qué punto estamos

El conocimiento colectivo empieza con distribución, pero depende de integración. Más personas pueden añadir observaciones, habilidades y perspectivas, aunque esas aportaciones no tienen efecto epistémico si son redundantes o quedan suprimidas. El primer problema de diseño es la independencia: un grupo se beneficia de estimaciones múltiples solo cuando sus miembros aportan información que no es simple copia mutua. Ahí la agregación puede ayudar o fallar.

2. La independencia vuelve informativa la agregación

Cuando los errores individuales tienen signos distintos, la agregación puede cancelar parte de ellos. El beneficio disminuye cuando las personas comparten fuentes, señales o incentivos porque los errores se correlacionan. Lorenz y colegas mostraron experimentalmente que la influencia social podía reducir la diversidad de estimaciones mientras aumentaba la confianza de los individuos.[2] La convergencia puede parecer mejora incluso cuando elimina la variación que la agregación necesitaba.

Afirmación C02. La influencia social puede reducir la diversidad de estimaciones y debilitar algunos beneficios de la agregación colectiva. [2]

En qué punto estamos

La agregación es poderosa cuando las contribuciones contienen errores parcialmente independientes. La influencia social puede reducir esa independencia antes de combinar estimaciones, creando confianza sin precisión equivalente. Un método colectivo debería distinguir, cuando sea posible, obtención de información y convergencia social. La pregunta siguiente es más amplia que promediar números: los grupos también difieren en las perspectivas de resolución que aportan a una tarea.

3. La diversidad puede añadir rutas de búsqueda

Hong y Page modelaron agentes que usan heurísticas distintas e identificaron condiciones bajo las que un grupo diverso de solucionadores capaces puede superar a un grupo formado solo por los mejores individuos.[1] El resultado es un teorema dentro de un modelo especificado, no una ley universal según la cual toda diversidad mejora cualquier tarea. Su valor es mostrar cómo la heterogeneidad puede crear rutas adicionales de búsqueda.

Afirmación C03. Heurísticas diversas de resolución de problemas pueden superar a grupos de solucionadores de gran capacidad bajo condiciones específicas del modelo. [1]

Afirmación C04. El resultado de Hong y Page depende de un modelo formal y no es un teorema universal sobre diversidad. [1]

En qué punto estamos

La diversidad puede contribuir creando rutas de búsqueda distintas, pero su valor depende de la tarea y de que los miembros puedan aportar sus diferencias. El resultado de Hong y Page es más fuerte como demostración de mecanismo que como conclusión universal. El grupo debe combinar perspectivas eficazmente. Los experimentos de resolución de problemas permiten probar si los grupos superan realmente a sus miembros más fuertes bajo condiciones controladas.

4. Algunos grupos superan a su mejor miembro

Laughlin, Hatch, Silver y Boh compararon individuos y grupos pequeños en problemas estructurados de letras y números. Grupos de tres, cuatro y cinco resolvieron los problemas en menos ensayos que los mejores individuos extraídos de números equivalentes de participantes, mientras que los grupos de dos no superaron al mejor de dos individuos.[3] La tarea tenía respuestas demostrablemente correctas y permitía intercambiar hipótesis y retroalimentación, por lo que el hallazgo muestra ventaja grupal bajo condiciones definidas, no superioridad universal de los grupos.

Afirmación C05. Los grupos de tres o más personas superaron a los mejores individuos en las tareas de letras a números estudiadas por Laughlin y colaboradores. [3]

Afirmación C06 - Síntesis declarada. Esa ventaja grupal es específica de la tarea y no establece una superioridad universal de los grupos. **Síntesis declarada:** [3] respalda directamente la ventaja en las tareas de letras a números estudiadas; la cláusula que niega una superioridad universal fuera de esas condiciones es la síntesis.

En qué punto estamos

Los grupos pueden superar a su mejor miembro cuando la tarea tiene estructura corregible y la comunicación permite acumular progresos parciales. Eso establece una ganancia colectiva real sin implicar que todas las tareas funcionen igual. El mecanismo dirige la atención a la calidad de interacción: cómo se turnan los miembros, responden a propuestas y mantienen información útil en juego. Aquí la investigación sobre inteligencia colectiva pasa de composición a proceso.

5. La calidad de interacción forma parte de la inteligencia colectiva

Woolley y colegas describieron un factor general de inteligencia colectiva en una batería de tareas grupales. El rendimiento se asoció con factores como sensibilidad social y reparto más igualitario de los turnos de conversación, en vez de quedar explicado simplemente por la inteligencia individual máxima o media de los miembros.[4] El estudio muestra que patrones de interacción pueden predecir rendimiento en su contexto experimental; no establece un rasgo fijo de todo grupo.

Afirmación C07. El rendimiento de grupos de laboratorio se ha asociado con sensibilidad social y con un reparto más igualitario de los turnos de conversación. [4]

En qué punto estamos

El rendimiento colectivo depende del proceso además de la capacidad individual. Trabajos experimentales vinculan mejor desempeño con patrones de interacción que permiten circular la información en vez de concentrarla en una voz. La comunicación se convierte así en mecanismo epistémico. Pero también puede distorsionar. El siguiente problema es cómo conservar el disenso el tiempo suficiente para distinguir información minoritaria útil de ruido.

6. El disenso puede transportar información

Un miembro discrepante puede estar equivocado, pero el desacuerdo también puede mostrar que el grupo no ha integrado toda la información disponible. Si el consenso se premia demasiado pronto, los miembros pueden suprimir observaciones inciertas o adaptar estimaciones a la mayoría visible. Estimaciones anónimas, juicios preliminares independientes e informes minoritarios explícitos intentan conservar información hasta poder evaluarla en vez de borrarla por convergencia social.[2]

Afirmación C08 - Síntesis declarada. El disenso puede conservar información que una convergencia prematura borraría. **Síntesis declarada:** [2] respalda la pérdida de diversidad bajo influencia social; tratar el disenso como mecanismo que puede conservar información que de otro modo se borraría es la síntesis.

En qué punto estamos

El disenso tiene valor epistémico cuando conserva información que los mecanismos de consenso borrarían. Las instituciones pueden protegerlo obteniendo juicios independientes antes de discutir o registrando explícitamente razones minoritarias. La meta no es desacuerdo permanente, sino convergencia informativa. Eso plantea la siguiente pregunta: el conocimiento colectivo suele depender de especialización, donde ningún miembro puede verificar por sí solo toda la cadena.

7. La especialización crea conocimiento distribuido

El análisis de Hutchins sobre navegación describe trabajo cognitivo distribuido entre especialistas, procedimientos y artefactos.[5] Las comunidades científicas funcionan de forma similar: especialistas en instrumentos, estadística, campo y teoría pueden aportar partes distintas de una cadena evidencial. La especialización aumenta capacidad pero crea dependencias de confianza, de modo que estándares, documentación y controles en las interfaces pasan a formar parte de la infraestructura de evidencia.

Afirmación C09. La cognición colectiva especializada puede distribuirse entre personas, artefactos y procedimientos. [5]

Afirmación C10 - Síntesis declarada. El conocimiento colectivo crea dependencias de confianza en las interfaces entre funciones especializadas. **Síntesis declarada:** [5] respalda cognición distribuida entre especialistas, procedimientos y artefactos; afirmar que esa distribución crea dependencias de confianza en las interfaces entre funciones es la síntesis.

En qué punto estamos

La especialización permite que los grupos sepan cosas que ningún miembro podría mantener solo, pero convierte las interfaces en infraestructura epistémica. La confianza debe ligarse a roles, registros y controles, no a la ficción de que cada participante verifica todo. Esta arquitectura distribuida conecta KNO-023 con KNO-021 y KNO-022. La última sección pregunta cómo combinar experiencia sin convertir prestigio en evidencia.

8. La experiencia debe ponderarse sin convertirse en autoridad

Ponderar a todos por igual no siempre es apropiado porque un especialista puede poseer mediciones o conocimiento técnico que otros no tienen. Sin embargo, ponderar por experiencia es distinto de aceptar estatus como prueba. Un proceso colectivo defendible pregunta qué competencia es relevante para el subproblema y combina autoridad específica de dominio con razones explícitas, disenso documentado, estimaciones independientes cuando son útiles y oportunidades de revisión.

Afirmación C11 - Síntesis declarada. La experiencia debe ponderarse por su relevancia para el subproblema y no tratarse como prueba basada en estatus. **Síntesis declarada:** no se asigna ninguna fuente como apoyo directo; tanto la regla de ponderar por relevancia como el rechazo del estatus como prueba son síntesis metodológica explícita.

Afirmación C12 - Síntesis declarada. La inteligencia colectiva es una organización condicional que conserva e integra información distribuida. **Síntesis declarada:** [1-5] respaldan mecanismos distintos de diversidad, independencia, interacción y cognición distribuida; definir la inteligencia colectiva como su organización condicional es la síntesis.

En qué punto estamos

Los grupos pueden superar a individuos combinando estimaciones independientes, estrategias diversas, experiencia complementaria y discusión correctora. Los mismos mecanismos sociales pueden crear error correlacionado, supresión y certeza prematura. La inteligencia colectiva es organización condicional, no tamaño de multitud. KNO-020 posee cómo se evalúa el consenso; KNO-023 posee cómo los grupos humanos distribuyen, conservan y combinan información antes de alcanzarlo.

Lo que sabemos y cómo lo sabemos

- Los grupos pueden superar a su mejor miembro en algunas tareas intelectivas estructuradas.[3]
- Heurísticas diversas pueden mejorar la búsqueda grupal bajo supuestos de modelo especificados.[1]

- La influencia social puede reducir diversidad de estimaciones y debilitar algunos beneficios de agregación.[2]
- Patrones de interacción como el reparto de turnos se asocian con rendimiento colectivo en grupos de laboratorio.[4]

Lo que todavía no sabemos

- Hasta qué punto las medidas de inteligencia colectiva de laboratorio generalizan a instituciones duraderas y equipos expertos de alto riesgo.
- Qué dimensiones de diversidad importan para qué tipos de problemas científicos e ingenieriles.
- Cómo ponderar experiencia sin suprimir evidencia minoritaria informativa.
- Cómo cambian la comunicación digital y la mediación por IA la independencia, conformidad y combinación de información en grupos humanos.

Qué podría hacernos cambiar de opinión

- Evidencia robusta de que la estructura de interacción no añade valor predictivo controlando capacidad individual en tareas realistas.
- Experimentos que muestren que la influencia social mejora fiablemente la precisión pese a reducir independencia en dominios amplios.
- Estudios de campo donde diversidad de perspectivas relevantes perjudique sistemáticamente el rendimiento aun con integración eficaz.
- Evidencia institucional de que mecanismos explícitos de conservación del disenso reducen en vez de mejorar detección de errores en decisiones comparables.

Si solo recuerdas una cosa

Los grupos saben más que los individuos solo cuando su estructura conserva y combina información distribuida mejor de lo que conformidad, sesgo compartido o consenso prematuro la destruyen.

Artículos relacionados

- KNO-009 · Archivos, bibliotecas y memoria colectiva
- KNO-020 · Revisión, consenso y desacuerdo
- KNO-021 · Ciencia abierta y reproducibilidad

Referencias

[1] Lu Hong; Scott E. Page. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101(46), 16385–16389 (2004). DOI 10.1073/pnas.0403723101. <https://doi.org/10.1073/pnas.0403723101> [2] Jan Lorenz; Heiko Rauhut; Frank Schweitzer; Dirk Helbing. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* 108(22), 9020–9025 (2011). DOI 10.1073/pnas.1008636108. <https://doi.org/10.1073/pnas.1008636108> [3] Patrick R. Laughlin; Erin C. Hatch;

Jonathan S. Silver; Lee Boh. Groups perform better than the best individuals on letters-to-numbers problems: effects of group size. *Journal of Personality and Social Psychology* 90(4), 644–651 (2006). DOI 10.1037/0022-3514.90.4.644. <https://doi.org/10.1037/0022-3514.90.4.644> [4] Anita Williams Woolley; Christopher F. Chabris; Alex Pentland; Nada Hashmi; Thomas W. Malone. Evidence for a collective intelligence factor in the performance of human groups. *Science* 330(6004), 686–688 (2010). DOI 10.1126/science.1193147. <https://doi.org/10.1126/science.1193147> [5] Edwin Hutchins. *Cognition in the Wild*. The MIT Press (1995). DOI 10.7551/mitpress/1881.001.0001. <https://doi.org/10.7551/mitpress/1881.001.0001>

KNO-024 · Conocimiento perdido y sesgo de preservación

KNO · Formas de conocer

Juan Seguro

v0.3 · ES

Todos los derechos reservados

Resumen del artículo

Lo que sobrevive del pasado no es una muestra aleatoria de lo que existió. Los huesos se degradan, los manuscritos arden, las instituciones descartan registros, los formatos dejan de ser legibles y los conjuntos de datos desaparecen cuando fallan dispositivos o direcciones de contacto. El archivo superviviente queda así moldeado por procesos de conservación, selección y destrucción antes de que historiador o científico alguno lo interprete. La arqueología formalizó este problema mediante procesos de formación que separan la actividad pasada del registro material recuperado después.[1] La paleontología estudia igualmente cómo la conservación filtra organismos y ambientes.[2,3] Los archivos añaden decisiones humanas sobre qué se crea, guarda, describe y hace accesible.[4] La información digital elimina algunas restricciones físicas, pero introduce dependencia de formatos, software, migración y custodia activa.[5,6] Este artículo pregunta cómo deben cambiar las inferencias cuando la ausencia puede reflejar no conservación y no inexistencia, y cómo los sistemas de preservación pueden hacer visibles sus propios efectos selectivos.

En 60 segundos

Los registros sobreviven a través de filtros. En arqueología, los objetos pasan del uso activo al contexto arqueológico mediante procesos culturales y naturales de formación, de modo que lo excavado ya está transformado.[1] En paleontología, meteorización, transporte, enterramiento y mineralización hacen que unos organismos y ambientes tengan mucha más probabilidad que otros de entrar en el registro fósil.[2,3] Los archivos añaden otra capa: personas e instituciones deciden qué documentar, conservar, clasificar y permitir consultar.[4] Los medios digitales no eliminan la pérdida. Vines y colegas hallaron que la disponibilidad de conjuntos de datos disminuía fuertemente con la antigüedad del artículo en su muestra, ilustrando la fragilidad de datos dependientes de custodios individuales.[5] La preservación digital exige migración, comprobación de integridad, metadatos y continuidad institucional, no mera copia.[6] La lección metodológica es simple pero exigente: la ausencia en un registro superviviente solo es evidencia después de considerar los procesos que pudieron eliminar, ocultar o nunca registrar el objeto.

Por qué importa esta pregunta

Razonamos constantemente a partir de supervivientes: fósiles, cuadernos de laboratorio, expedientes administrativos, yacimientos, fotografías, conjuntos de datos y código fuente. Como esos supervivientes son visibles, pueden parecer el pasado mismo. No lo son. Todo registro tiene una historia de supervivencia, y esa historia puede estar correlacionada con la propiedad que queremos inferir. Los huesos grandes se conservan de forma distinta a tejidos frágiles; los archivos oficiales, de

forma distinta al habla informal; los repositorios mantenidos, de forma distinta a discos olvidados. El sesgo de conservación importa cuando el mecanismo de supervivencia puede distorsionar frecuencia, diversidad o cronología aparentes.

1. El registro es producido por la supervivencia

La pregunta Si la mayoría de las huellas desaparece antes de poder inspeccionarse, ¿qué muestrea realmente el registro superviviente? La respuesta depende no solo de lo que ocurrió en el pasado, sino también de qué huellas se crearon, conservaron, recopilaron, describieron y quedaron disponibles para observadores posteriores.

Un acontecimiento pasado y una huella superviviente no son el mismo objeto. Entre ambos actúan procesos que crean, transforman, mueven, conservan y destruyen registros. La distinción de Schiffer entre contexto sistémico y arqueológico convirtió los procesos de formación en parte central de la inferencia: los objetos recuperados tienen historias posteriores al uso que afectan ubicación y asociación.[1] Antes de tratar una huella como muestra del pasado, hay que preguntar qué hizo probable su supervivencia.

Afirmación C01. La evidencia arqueológica está moldeada por procesos de formación entre el uso activo y su recuperación posterior. [1]

En qué punto estamos

El registro superviviente se genera dos veces: primero por los acontecimientos que crean huellas y después por los procesos que conservan unas y eliminan otras. La inferencia debe modelar ambos. La conservación no es una molestia añadida tras observar; forma parte del propio proceso de observación. La sección siguiente muestra cuánto importa esto en el registro físico de organismos y ambientes.

2. La tafonomía filtra el pasado biológico

Behrensmeyer calibró etapas de meteorización ósea con cadáveres de edad conocida y mostró que la meteorización depende del tiempo y de condiciones locales.[2] Kidwell y Flessa revisan variaciones en completitud, resolución temporal y fidelidad ecológica del registro fósil.[3] La abundancia fósil no equivale por tanto a abundancia pasada: partes duras, enterramiento rápido y ambiente deposicional alteran la probabilidad de conservación. La tafonomía trata ese filtro como evidencia y no como ruido.

Afirmación C02. La meteorización ósea varía con el tiempo y las condiciones ambientales y puede informar la interpretación tafonómica. [2]

Afirmación C03. El registro fósil difiere en completitud y en resolución temporal y ecológica según el contexto. [3]

En qué punto estamos

Los fósiles son observaciones filtradas. Meteorización, transporte, enterramiento y mineralización cambian qué organismos, partes corporales e intervalos temporales se vuelven visibles. La tafonomía no vuelve inútil el registro fósil; convierte el proceso de conservación en parte de la evidencia. El mismo principio se aplica a registros humanos, con una complicación adicional: personas e instituciones deciden intencionadamente qué documentar y conservar.

3. Los archivos se construyen, no solo se heredan

Schwartz y Cook sostienen que archivos y registros participan en la construcción de la memoria social en vez de funcionar como contenedores neutrales de hechos.[4] Decisiones sobre creación, valoración, descripción, acceso y custodia modelan lo que usuarios posteriores pueden encontrar. Un registro ausente puede significar que algo no ocurrió, pero también que no se creó, se destruyó, quedó fuera de adquisición, fue mal descrito o tuvo acceso restringido.

Afirmación C04. Los archivos y registros están moldeados por decisiones humanas sobre creación, valoración, descripción y acceso. [4]

En qué punto estamos

Los archivos humanos añaden selección intencional a la supervivencia física. Los registros pueden faltar porque nadie los creó, porque una institución no los conservó, porque la descripción los volvió difíciles de localizar o porque el acceso quedó restringido. La procedencia archivística incluye por tanto custodia y política. La siguiente cuestión es la asimetría: los sistemas oficiales suelen conservar unas voces y actividades más sistemáticamente que otras.

4. La ausencia suele estar estructurada

El sesgo de conservación es más peligroso cuando la ausencia se correlaciona con la pregunta estudiada. Piedra y metal duraderos sobreviven de forma distinta a textiles, madera o habla; las instituciones formales documentan de forma distinta a comunidades informales; las crisis pueden dejar más papeles que el éxito rutinario. Las preguntas clave son qué categorías podían crear huellas, cuáles tenían probabilidad de sobrevivir y cuáles de ser recogidas y descubiertas.

Afirmación C05 - Síntesis declarada. La ausencia puede estar estructurada cuando la probabilidad de supervivencia depende de la propiedad que se estudia. **Síntesis declarada:** [1-4] respaldan procesos selectivos de formación, supervivencia y archivo; describir la ausencia como estructurada cuando la supervivencia depende de la propiedad estudiada es la síntesis.

Afirmación C06 - Síntesis declarada. La ausencia en un archivo no es automáticamente evidencia de que un acontecimiento no ocurrió. **Síntesis declarada:** [1,4] respaldan varias vías por las que un registro puede no formarse o no sobrevivir; inferir que la ausencia archivística no equivale automáticamente a ausencia del acontecimiento es la síntesis.

En qué punto estamos

El sesgo de conservación es más fuerte cuando el mecanismo de pérdida está relacionado con el fenómeno estudiado. Entonces el registro superviviente puede exagerar materiales duraderos, oficiales, excepcionales o valorados institucionalmente. Interpretar correctamente exige un modelo de quién podía crear registros y cuáles podían sobrevivir. Los medios digitales parecen eliminar muchos filtros físicos, pero introducen dependencias y modos de fallo propios.

5. Digital no significa permanente

Vines y colegas solicitaron conjuntos de datos de 516 artículos biológicos con edades entre dos y veintidós años y hallaron que, entre los casos donde los autores informaron del estado de los datos, las probabilidades en favor de que un conjunto siguiera existiendo disminuyeron un 17 % por año en

su muestra.[5] Direcciones de contacto rotas y almacenamiento obsoleto fueron obstáculos importantes. La estimación es específica de la muestra, pero demuestra que los datos digitales pueden desaparecer rápidamente sin preservación institucional.

Afirmación C07. En la muestra de Vines y colaboradores, la disponibilidad de datos disminuyó con fuerza al aumentar la antigüedad del artículo. [5]

Afirmación C08. La preservación digital requiere una custodia activa que va más allá de la simple copia. [5,6]

En qué punto estamos

La información digital puede copiarse con bajo coste, pero la persistencia sigue requiriendo custodia activa. Los datos pueden desaparecer por dispositivos obsoletos, enlaces rotos, credenciales perdidas, deriva de formatos o pérdida de contexto. El estudio de Vines muestra una forma empírica de ese deterioro en datos de investigación. La preservación pasa así de proteger un objeto físico a mantener una cadena interpretable de archivos, software, metadatos e instituciones.

6. Digitalizar no equivale a preservar digitalmente

Conway distingue digitalización para conservación de preservación digital como actividad sociotécnica. [6] Escanear puede reducir manipulación del original y ampliar acceso, pero el escaneo se convierte en un objeto digital que requiere gestión. Resolución, color, compresión, metadatos y formato determinan qué captura el sustituto, mientras migración y validación determinan si sigue siendo utilizable.

Afirmación C09. La digitalización y la preservación digital son actividades distintas. [6]

Afirmación C10 - Síntesis declarada. Un sustituto digital conserva solo lo que codifican sus decisiones de captura y metadatos. **Síntesis declarada:** [6] respalda distinguir decisiones de digitalización de preservación digital; la cláusula de que el sustituto conserva solo lo que codifican captura y metadatos es la síntesis.

En qué punto estamos

La digitalización crea un nuevo objeto de conservación en vez de terminar el trabajo de preservar. Lo capturado depende de decisiones de conversión, y los archivos resultantes requieren gestión continua. Un sustituto digital puede mejorar acceso y omitir rasgos del original o quedar obsoleto. El problema siguiente es que los sistemas de conservación deben hacer visibles sus propias selecciones y transformaciones para que usuarios futuros juzguen qué sobrevivió.

7. La conservación necesita procedencia de pérdidas y cambios

Un archivo útil puede registrar no solo lo que contiene sino, cuando es posible, qué fue transformado, migrado, retirado o se sabe que falta. Las sumas de comprobación documentan integridad de bytes; los registros de migración, cambios de formato; las políticas de colección, límites de selección. Estos registros no reconstruyen todo lo perdido, pero vuelven auditables algunos procesos de conservación y ayudan a distinguir evidencia original de transformaciones posteriores.

Afirmación C11 - Síntesis declarada. Los sistemas de preservación pueden mejorar la auditabilidad registrando migraciones, integridad e historial de selección. **Síntesis declarada:** [6] respalda la necesidad de gestionar objetos digitales a través de acciones de preservación; afirmar que registrar migraciones, integridad e historial de selección mejora la auditabilidad es la síntesis.

En qué punto estamos

Los sistemas de conservación pueden reducir sesgo no solo guardando más, sino documentando los filtros que aplican. Registros de integridad, historiales de migración y políticas de colección exponen partes del proceso de supervivencia al escrutinio posterior. No recuperan lo que nunca se registró, pero evitan que una transformación silenciosa se presente como evidencia original. La sección final pregunta cómo formular afirmaciones cuando el filtro de conservación solo se conoce parcialmente.

8. La ausencia requiere un modelo de conservación

Una ausencia es evidencia más fuerte cuando el sistema habría tenido alta probabilidad de crear, conservar y exponer el registro pertinente si el acontecimiento hubiera ocurrido. Es más débil cuando alguna de esas probabilidades es baja o desconocida. Arqueólogos modelan procesos de formación, paleontólogos estudian tafonomía, archiveros examinan valoración y custodia, y la preservación digital estudia formatos y custodia. Dominios distintos expresan la misma disciplina: interpretar el registro mediante el proceso que lo hizo observable.

Afirmación C12 - Síntesis declarada. La ausencia es evidencia más fuerte cuando la creación, preservación y posibilidad de descubrimiento habrían sido todas probables. **Síntesis declarada:** [1-6] respaldan filtros distintos de creación, supervivencia, archivo y preservación digital; combinarlos en este criterio sobre cuándo la ausencia constituye evidencia más fuerte es la síntesis.

En qué punto estamos

El conocimiento perdido no es solo contenido ausente; es incertidumbre sobre el proceso de filtrado que produjo lo que queda. La ausencia se vuelve informativa cuando creación, conservación y descubribilidad habrían sido probables. Cuando esas condiciones son débiles, la confianza debe bajar. KNO-024 cierra la secuencia Formas de conocer convirtiendo la conservación en parte de la evidencia: lo que sobrevive limita el conocimiento solo mediante un modelo de cómo sobrevivió.

Lo que sabemos y cómo lo sabemos

- Los registros arqueológicos están moldeados por procesos de formación entre uso y recuperación.[1]
- La meteorización ósea y otros procesos tafonómicos afectan qué huellas biológicas sobreviven y cómo se interpretan.[2,3]
- Los archivos incorporan decisiones de creación, valoración, descripción y acceso en vez de ser contenedores neutrales.[4]
- Los conjuntos de datos pueden dejar de estar disponibles con el tiempo cuando la conservación depende de custodios individuales.[5]

Lo que todavía no sabemos

- Cuánto material no documentado se ha perdido ya de ecosistemas contemporáneos de investigación digital.
- Cómo varían las probabilidades de conservación entre comunidades infrarrepresentadas y sistemas informales de registro.
- Qué estrategias de preservación digital seguirán siendo interpretables tras cambios tecnológicos de siglos.
- Cómo cuantificar mejor el sesgo de conservación cuando creación, pérdida y descubrimiento son inciertos.

Qué podría hacernos cambiar de opinión

- Registros empíricos que muestren que clases importantes de evidencia supuestamente frágil se conservan al mismo ritmo que evidencia duradera en contextos relevantes.
- Estudios archivísticos que demuestren que valoración y acceso no afectan de forma medible inferencias históricas posteriores en colecciones comparables.
- Estudios longitudinales digitales que muestren pérdida insignificante de datos sin custodia institucional.
- Nuevos métodos de conservación que vuelvan irrelevante a largo plazo la obsolescencia de formatos, software y metadatos.

Si solo recuerdas una cosa

Un registro superviviente es evidencia sobre el pasado y sobre los procesos que seleccionaron qué podía sobrevivir; la ausencia solo se interpreta considerando ambos.

Artículos relacionados

- KNO-009 · Archivos, bibliotecas y memoria colectiva
- KNO-019 · Datos, tablas y registros
- KNO-021 · Ciencia abierta y reproducibilidad

Referencias

[1] Michael B. Schiffer. Archaeological Context and Systemic Context. *American Antiquity* 37(2), 156–165 (1972). DOI 10.2307/278203. <https://doi.org/10.2307/278203> [2] Anna K. Behrensmeyer. Taphonomic and ecologic information from bone weathering. *Paleobiology* 4(2), 150–162 (1978). DOI 10.1017/S0094837300005820. <https://doi.org/10.1017/S0094837300005820> [3] Susan M. Kidwell; Karl W. Flessa. The quality of the fossil record: populations, species, and communities. *Annual Review of Earth and Planetary Sciences* 24, 433–464 (1996). DOI 10.1146/annurev.earth.24.1.433. <https://doi.org/10.1146/annurev.earth.24.1.433> [4] Joan M. Schwartz; Terry Cook. Archives, Records, and Power: The Making of Modern Memory. *Archival Science* 2(1–2), 1–19 (2002). DOI 10.1007/BF02435628. <https://doi.org/10.1007/BF02435628> [5] Timothy H. Vines; Arianne Y.K. Albert; Rose L. Andrew; Florence Débarre; Dan G. Bock; Michelle T. Franklin; Kimberly J. Gilbert; Jean-Sébastien Moore; Sébastien

Renaut; Diana J. Rennison. The availability of research data declines rapidly with article age. *Current Biology* 24(1), 94–97 (2014). DOI 10.1016/j.cub.2013.11.014. <https://doi.org/10.1016/j.cub.2013.11.014>

[6] Paul Conway. Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas. *The Library Quarterly* 80(1), 61–79 (2010). DOI 10.1086/648463. <https://doi.org/10.1086/648463>