

El problema oculto del observador: puntos finales subjetivos, expectativa y lector no cegado

Juan Segurola

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Segurola. Todos los derechos reservados.

Correspondencia: info@3Dresyns.com · www.3Dresyns.com

HEB-006 · Artículo técnico independiente · Versión v0.2

DOI: 10.5281/zenodo.22915113

Resumen

Algunos puntos finales experimentales quedan completados por un observador humano. Se juzga si una grieta está presente o ausente, se califica una superficie como aceptable, se declara una coincidencia de color, se asigna una probeta a una categoría, se segmenta una imagen, se clasifica un modo de fallo o se acepta o rechaza una traza ambigua. En estos casos, el observador no es externo al sistema de medición. La percepción, la formación, la expectativa, la información disponible en el momento del juicio, las definiciones de categoría y las reglas de parada pueden influir en el resultado registrado. Este artículo desarrolla el problema oculto del observador como un fallo metodológico general en el que decisiones subjetivas o semisubjetivas se tratan como si fueran propiedades directas de la probeta. La igualdad histórica de colores ofrece un contraejemplo constructivo: las mediciones de Maxwell trataron la igualdad perceptiva como una relación empírica que podía repetirse y compararse entre observadores, en lugar de confundir igualdad perceptiva con identidad física. Ejemplos contemporáneos de ingeniería muestran los riesgos complementarios: «fácil de retirar» no es evidencia transferible y la inspección visual de geometrías ocultas no es adecuada salvo que se haya validado. El artículo distingue los puntos finales humanos legítimos de la dependencia no controlada del observador y propone una arquitectura práctica basada en criterios predefinidos, cegamiento cuando sea viable, aleatorización, múltiples observadores, análisis del desacuerdo, ejemplos de referencia, lecturas repetidas, anclaje objetivo y conservación de los casos ambiguos. La conclusión central es que un juicio humano puede ser un componente válido de la medición, pero solo cuando el papel del observador se hace explícito, acotado y comprobable.

Palabras clave: observador; punto final subjetivo; cegamiento; inspección visual; variación entre evaluadores; expectativa; clasificación; igualdad de color; sistema de medición; reproducibilidad; inferencia científica

1. Propósito y límite de evidencia

El lenguaje científico a menudo oculta al observador dentro de verbos que suenan objetivos: una superficie «parecía» uniforme, una probeta «mostró» fallo, una imagen «contenía» un defecto, un color «coincidía», una traza era «aceptable» o una muestra era «claramente» diferente. A veces estas afirmaciones describen observaciones robustas. Otras veces comprimen una decisión humana no examinada.

El objetivo de este artículo no es eliminar el juicio humano. Muchos puntos finales científicamente importantes son perceptivos, morfológicos o clasificatorios y no pueden reducirse a una única salida instrumental sin perder información. El objetivo es tratar al observador como un componente del sistema de medición siempre que el resultado registrado dependa de la interpretación humana.

HEB-001 establece que el sistema de medición incluye al operador, además del instrumento, la probeta, la referencia, el entorno, el protocolo y el tratamiento estadístico [1]. HEB-006 desarrolla las consecuencias epistémicas de esa inclusión.

La proposición central es:

Cuando un juicio humano completa la medición, las condiciones que conforman ese juicio forman parte del método.

2. Observación no equivale a subjetividad sin restricciones

La observación humana puede ser muy informativa. La microscopía, la patología, el análisis de superficies de fractura, la igualdad de colores, la clasificación de defectos y muchas inspecciones visuales dependen de observadores entrenados. El problema metodológico comienza cuando un juicio se trata como autoválido por el mero hecho de haberlo emitido una persona experimentada.

Conviene distinguir entre acceso perceptivo e interpretación inferencial. Un observador puede detectar de forma fiable una discontinuidad visible y, al mismo tiempo, no estar seguro de su causa. Un lector puede identificar una

característica espectral y discrepar sobre el mecanismo que la produjo. Una igualación de color puede ser reproducible aunque los dos espectros sean físicamente distintos. En ciencia del color, los estímulos físicamente diferentes que producen la misma igualación de color bajo condiciones de observación definidas se denominan metameros; el fenómeno se denomina metamerismo.

El problema del observador contiene, por tanto, al menos tres capas: detección, clasificación y explicación. No deben colapsarse en una sola.

3. El trabajo de Maxwell sobre color: el observador como parte del aparato

El proyecto Maxwell de este corpus resalta un caso histórico útil. Los experimentos de color de Maxwell conectaban estímulos físicos controlados con igualaciones perceptivas realizadas por observadores reales. Dos espectros físicamente diferentes podían juzgarse perceptivamente iguales sin ser físicamente idénticos. La igualdad era empírica: podía reproducirse, cuantificarse y compararse entre observadores [2,5].

La importancia es metodológica. El observador no se trataba como una molestia que hubiera que eliminar del experimento. La percepción era el punto final, por lo que el observador pertenecía necesariamente al aparato. El avance científico consistió en restringir esa relación dependiente del observador mediante estímulos controlados, proporciones ajustables y comparación entre personas.

Este ejemplo también muestra por qué la dependencia del observador no vuelve irrelevante un resultado. Un resultado puede depender del observador y seguir siendo reproducible, acotado y científicamente potente. El requisito es definir la respuesta de quién, bajo qué condiciones, ante qué estímulo y con qué repetibilidad.

4. La expectativa puede entrar antes de registrar el juicio

Un observador puede saber que probeta es el control, que formulación se espera que funcione mejor, que imagen pertenece al proceso preferido o que punto de datos haría encajar un modelo. Esa información puede influir en la atención, la selección de umbrales, la categorización y la disposición a repetir o rechazar una observación.

Esto no requiere mala conducta. Es una propiedad estructural de una decisión no cegada. Si el observador dispone de información correlacionada con el resultado esperado, el sistema de medición incluye esa vía de información.

El cegamiento no es, por tanto, un ritual. Es un método para eliminar un canal de información irrelevante para el punto final que se está juzgando. ER-209 exige aleatorización y duplicados ciegos en trabajos interlaboratorio precisamente porque, de otro modo, la información de identidad puede quedar entrelazada con la interpretación y el tratamiento de datos [3].

5. Caso de estudio I: «fácil de retirar» no es una medición

ER-250 ofrece un ejemplo contemporáneo de ingeniería. Establece qué, cuando el riesgo de daño durante la retirada de soportes es significativo, la fuerza o el par de retirada deben medirse con probetas representativas y que una descripción subjetiva como «fácil de retirar» no constituye evidencia transferible [4].

La expresión «fácil de retirar» comprime varias variables: fuerza del operador, herramienta elegida, ángulo de acceso, secuencia, velocidad, expectativa, experiencia previa, geometría de la pieza y umbral interno del observador para distinguir entre fácil y difícil. Puede ser útil como nota exploratoria, pero no es un mensurando estable de ingeniería hasta que el juicio se ancla a condiciones controladas o a una respuesta objetiva.

La lección general no es que los descriptores subjetivos estén prohibidos. Es que un adjetivo no calibrado no debe heredar silenciosamente el estatus probatorio de una medición.

6. Caso de estudio II: la inspección visual está limitada por lo que puede verse

ER-250 también establece que la inspección visual por sí sola no es adecuada para geometrías ocultas salvo que se valide, y pide una sensibilidad de detección definida mediante defectos sembrados u otro método acotado [4]. Esto ilustra un fallo diferente del observador: el punto final está limitado no solo por el juicio, sino también por el acceso.

Un inspector puede ser perfectamente consistente y aun así pasar por alto defectos ocultos, por debajo de la resolución, mal iluminados o confundidos por la textura superficial. Repetir la misma inspección no recupera información que nunca llegó al observador.

La validación del observador requiere, por tanto, tanto fiabilidad perceptiva como disponibilidad del estímulo. Un juicio reproducible sobre una característica inaccesible sigue siendo una medición inválida de esa característica.

7. Las categorías pueden crear una certeza artificial

Muchos puntos finales mediados por observador son categoricos: pasa/fallá, grieta/sin grieta, aceptable/no aceptable, fase A/fase B, transparente/opaco, curado/no curado. La realidad subyacente puede ser continua o multidimensional. Por tanto, un límite de categoría puede crear una transición registrada abrupta donde el sistema físico cambia gradualmente.

El método debe definir el umbral antes de juzgar los datos. Si el límite se desplaza después de ver los resultados, el observador pasa a formar parte del proceso de ajuste del modelo. Los casos fronterizos deben conservarse en lugar de forzarse a una categoría solo para completar una tabla.

Una práctica útil es registrar tanto la categoría como la evidencia continua o descriptiva subyacente cuando este es disponible. Esto permite a lectores posteriores comprobar si la conclusión depende del umbral.

8. La formación mejora la consistencia, pero no demuestra validez

La formación puede alinear a los observadores en torno a definiciones y ejemplos compartidos. Imágenes de referencia, atlas de defectos, patrones de color y ejemplos resueltos pueden reducir variación evitable. Pero la concordancia entre observadores entrenados todavía puede ser errónea si el criterio compartido no está vinculado al punto final físico o funcional pretendido.

La concordancia entre observadores y la validez responden preguntas distintas. Una alta concordancia muestra que una regla puede aplicarse de forma consistente. La validez requiere evidencia de que la regla identifica aquello que la afirmación científica dice que identifica.

La misma distinción se aplica a clasificadores automatizados. Sustituir una persona por software puede mejorar la consistencia y conservar, al mismo tiempo, las mismas etiquetas sesgadas, características inaccesibles o una verdad de referencia inadecuada.

9. Múltiples observadores convierten el desacuerdo en datos

Cuando un punto final depende del juicio, un segundo observador no es simplemente un duplicado. El desacuerdo mide algo sobre la ambigüedad del punto final, la adecuación de la definición de categoría, la visibilidad de la característica o el estado de formación.

El caso de color de Maxwell demuestra el valor constructivo de comparar observadores: un segundo observador ayuda a separar una relación peculiar de un sistema visual de una relación reproducida en otro [2,5]. En experimentos modernos, la misma lógica sustenta lecturas independientes, adjudicación ciega e informe explícito del desacuerdo.

El consenso no debe borrar el registro original del desacuerdo. Si se requiere adjudicación, los juicios independientes iniciales y la regla utilizada para resolverlos deben seguir siendo recuperables.

10. La lectura repetida puede medir la repetibilidad del observador

Un observador puede evaluarse frente a sus propias decisiones previas. Volver a leer probetas o imágenes en orden aleatorio después de un intervalo permite estimar la repetibilidad intraobservador. Repetir solo los casos obvios no aporta información; el conjunto de desafío debe incluir casos fronterizos y negativos y positivos conocidos cuando puedan establecerse.

La lectura repetida es especialmente valiosa cuando el punto final se utiliza para formar a otras personas o para crear conjuntos de datos etiquetados. Si el observador de referencia no es consistente consigo mismo, las etiquetas no pueden tratarse como una verdad de referencia libre de ruido.

11. El anclaje objetivo debe complementar, no sustituir automáticamente, el juicio

Una medición objetiva puede anclar un punto final subjetivo cuando ambas abordan partes relacionadas de la misma pregunta: la fuerza de retirada medida puede acompañar a la facilidad percibida; la perfilometría puede acompañar al aspecto superficial; métricas de imagen pueden acompañar a las valoraciones visuales de defectos; mediciones espectrales pueden acompañar a la igualación de color.

No debe suponerse que el anclaje es una verdad superior si el punto final científico es genuinamente perceptivo. En la igualación de color, por ejemplo, la igualdad perceptiva es en sí misma el fenómeno de interés. Los espectros físicos explican y restringen el estímulo, pero no sustituyen el punto final perceptivo.

La pregunta de diseño es, por tanto: ¿qué partes del punto final dependen intrínsecamente del observador y cuáles pueden medirse de forma independiente para poner a prueba el juicio?

12. Una arquitectura práctica de control del observador

Una medición robusta mediada por observador puede diseñarse mediante una secuencia de controles: definir el punto final antes de observar; especificar la evidencia observable; estandarizar las condiciones de visualización o lectura; eliminar información de identidad irrelevante cuando sea posible; aleatorizar el orden; proporcionar ejemplos de referencia; incluir casos ambiguos y de desafío; obtener lecturas independientes cuando la consecuencia sea elevada; cuantificar concordancia y repetibilidad; conservar los desacuerdos; y anclar el juicio a evidencia física independiente cuando corresponda.

Estos controles deben ser proporcionales a la afirmación. Una nota exploratoria de baja consecuencia no requiere la misma arquitectura que un criterio de aceptación, un punto final clínico o una decisión de seguridad. Pero cuanto más trascendente sea la conclusión, menos defendible resulta un juicio de lector único no documentado.

13. Modos de fallo del observador

Entre los modos de fallo recurrentes están la fuga de expectativas, la identidad no cegada, umbrales inconsistentes, malas condiciones de iluminación o pantallá, fatiga, efectos de orden, características inaccesibles, compresión categórica, deriva de la formación, consenso sin lecturas independientes, repetición selectiva de casos dudosos y uso del mismo observador para generar y validar un conjunto de etiquetas de referencia.

Otro fallo aparece cuando la interpretación se registra como observación. «Grieta presente» y «fallo causado por fragilidad» no son la misma afirmación. La primera puede ser una clasificación visual; la segunda es una inferencia mecánica que requiere evidencia adicional.

El lenguaje más seguro nombra lo que el observador vio realmente antes de nombrar lo que la observación puede significar.

14. Relacion con el corpus Engineering Review

HEB-006 aborda la dependencia del observador como problema científico general. HEB-002 aborda límites conceptuales ocultos, HEB-003 aborda el razonamiento visual como vía desde la estructura hasta la predicción, HEB-004 aborda el desajuste proxy-mensurando y HEB-005 aborda la validez de la referencia. HEB-006 es distinto porque el observador o lector participa directamente en la detección, clasificación o interpretación. ER-209 conserva el diseño interlaboratorio, incluida la aleatorización y el cegamiento. ER-250 conserva la medición de retirada de soportes, los efectos del operador, la sensibilidad de inspección y la cualificación funcional [3,4]. Otros Engineering Reviews pueden contener puntos finales visuales o manuales específicos de aplicación, pero HEB-006 no sustituye sus requisitos de cualificación.

El límite es el mismo utilizado en toda la serie HEB: este artículo pregunta si una conclusión se deriva de una medición mediada por observador; el corpus ER pregunta como debe calificarse un punto final concreto de ingeniería.

15. Conclusiones

Un observador humano puede ser un instrumento científico indispensable. El error metodológico consiste en fingir que el observador está ausente cuando la percepción, la clasificación o la interpretación completan la medición.

Las principales conclusiones son:

1. Los puntos finales mediados por observador son legítimos cuando el papel del observador es explícito y está controlado.
2. Detección, clasificación e interpretación causal son etapas distintas y no deben colapsarse.
3. El cegamiento elimina canales de información irrelevantes que, de otro modo, pueden pasar a formar parte del sistema de medición.
4. Los descriptores subjetivos pueden apoyar la exploración, pero no deben tratarse como evidencia cuantitativa transferible sin anclaje.
5. La inspección visual está limitada por la visibilidad de la característica y la sensibilidad de detección, además del juicio del observador.
6. La concordancia entre observadores mide consistencia, no necesariamente validez.
7. El desacuerdo es informativo y debe conservarse en lugar de ocultarse mediante consenso forzado.
8. La lectura repetida puede estimar la fiabilidad intraobservador y revelar reglas de clasificación inestables.
9. Las mediciones objetivas pueden anclar el juicio, pero no deben borrar puntos finales que son intrínsecamente perceptivos.
10. El método debe registrar lo observado antes de escalar hacia lo que se afirma que significa esa observación.

El observador solo se vuelve peligroso para el experimento cuando el experimento olvida que el observador estaba allí.

Referencias

- [1] Seguro, J. The Hidden Error Budget: Measurement-System Failure, False Precision and Experimental Design in Laboratory Science. HEB-001. 3Dresyns Independent Technical Articles (2026). DOI: 10.5281/zenodo.22862642.
- [2] Seguro, J. Maxwell: Thinking the Invisible - James Clerk Maxwell and the Art of Reasoning Beyond What Could Be Seen. Versión 1.3, 11 September 2026. DOI: 10.5281/zenodo.22714429.
- [3] Seguro, J. Interlaboratory Round-Robin Design for Vat-Photopolymer Materials: Specimen state, protocol transfer and reproducibility statistics. ER-209. 3Dresyns Engineering Publications (2026). Versión DOI: 10.5281/zenodo.22862020.
- [4] Seguro, J. Support-Removal Effects on Fine Features and Internal Geometries: Access, Load Transfer, Surface Damage and Verification. ER-250. 3Dresyns Engineering Publications (2026). Versión DOI: 10.5281/zenodo.22883403.
- [5] Maxwell, J. C. On the Theory of Compound Colours, and the Relations of the Colours of the Spectrum. Philosophical Transactions of the Royal Society of London 150, Part I (1860) 57-84.

Límite de conocimiento

Este artículo utiliza ejemplos relacionados con el observador ya documentados dentro del corpus controlado del proyecto. El ejemplo de Maxwell se limita al punto probatorio respaldado allí: las igualaciones perceptivas se midieron y compararon entre observadores, y la igualdad perceptiva no se equipara con identidad física. Los ejemplos de ingeniería se utilizan solo para ilustrar el control del observador y no sustituyen sus contextos originales de cualificación.

Limitaciones

El artículo es una síntesis metodológica, no una revisión cuantitativa de tamaños de efecto del sesgo del observador ni de estadísticas entre evaluadores. No se prescribe un número universal de observadores, un umbral de concordancia ni un procedimiento de cegamiento. Los controles apropiados dependen del punto final, la consecuencia, la verdad de referencia disponible y de si la respuesta dependiente del observador es en si misma el fenómeno de interés.