VISUAL REVIEW | Vol. 18, No. 6, 2026 | ISSN 2695-9631
International Visual Culture Review / Revista Internacional de Cultura Visual
https://doi.org/10.62161/revvisual.v18.6439

TEST VISUAL DE TURING Y ATRIBUCIÓN DE AUTORÍA
Indicios visuales, confianza y precisión en la identificación de imágenes generadas por IA en profesionales de las artes visuales

ANDRÉS TORRES-CARCELLER [1]
andrestorresb.edu
JÚLIA CASTELL-VILLANUEVA [1]
julia.castellb.edu
[1]Universidad de Barcelona, España

PALABRAS CLAVE
IA generativa
test visual de Turing
imagen
atribución de autoría
metacognición

RESUMEN
Este estudio adapta un test de Turing visual para examinar la atribución de autoría (humana vs IAGen) en artes visuales (N=28) con 10 estímulos equilibrados. Por imagen se registraron decisión, confianza (1–5), pesos de cinco familias de indicios y valoración estética antes y después de revelar la autoría. La precisión global fue moderada (57,1%), con falsos positivos en humanos de acabado digital y falsos negativos en registros abstractos o lineales. La confianza no calibró el acierto y predijo menor precisión; estructura/espacio e iluminación/color se asociaron a mayor acierto. Conocer la autoría devaluó más la IAGen.

Recibido: 03/ 07 / 2026
Aceptado: 05/ 10 / 2026

1. Introducción

En 1950, Alan Turing abordó la cuestión de la inteligencia desde una decisión metodológica: en lugar de discutir su esencia, propuso un modo de validación mediante una situación evaluable. Su propuesta consistía en comprobar si una máquina podía sostener una interacción textual bajo condiciones controladas sin ser identificada como tal. La fuerza de ese giro no reside en ofrecer una definición, sino en situar el problema en un umbral práctico: el punto en que las categorías disponibles dejan de orientar el juicio con seguridad. Trasladado a la cultura visual, el planteamiento adquiere una forma inmediata: qué sostiene la decisión cuando la autoría de una imagen ya no puede inferirse de manera fiable.

El desarrollo reciente de la Inteligencia Artificial Generativa (IAGen) ha intensificado un fenómeno que atraviesa hoy el conjunto de la cultura visual: la creciente dificultad para distinguir entre la producción humana y la algorítmica. Si bien los primeros trabajos se concentraron en la verosimilitud fotográfica y la síntesis de rostros —debido a su impacto social en forma de deepfakes (Korshunov y Marcel, 2018)—, el problema actual trasciende el fotorrealismo. Lo que aquí interesa es la capacidad de la IAGen para generar creaciones con competencia estilística y calidad artística en registros diversos, emulando convenciones y soluciones formales que el observador asocia habitualmente a prácticas autorales. En este terreno, la atribución se vuelve incierta porque no depende de detectar un error técnico, sino de decidir si la coherencia visual y la intencionalidad aparente remiten a un proceso humano o a una síntesis algorítmica capaz de simular con solvencia esos mismos rasgos. Esta convergencia entre los códigos visuales de los artistas y los patrones recombinados por los modelos generativos (Nightingale y Farid, 2022) introduce una zona de indeterminación perceptiva que la estética tradicional no logra resolver únicamente mediante la observación.

Desde una perspectiva educativa y cultural, esta indistinción plantea un problema de fondo. La imagen artística ha funcionado históricamente como un espacio privilegiado para atribuir intención, expresividad y experiencia. Cuando estas cualidades son simuladas de manera convincente por sistemas no humanos, se debilitan los criterios que sustentan el juicio sobre autoría, autenticidad y valor (Carceller, 2024). Sin embargo, la cuestión no debe formularse como una exigencia de identificación constante, sino como la constatación de un umbral de incertidumbre: aquel donde la lectura ya no permite inferir el origen de la imagen y obliga a reconsiderar el objeto de la evaluación.

En un escenario donde la IAGen se integra como herramienta creativa, el interés se desplaza hacia cómo se sostiene el juicio cuando el origen deja de funcionar como apoyo estable. En este marco, la adaptación del test de Turing al ámbito de las artes visuales emerge como un dispositivo analítico para examinar no solo la capacidad técnica de la IA, sino los procesos cognitivos mediante los cuales los sujetos emiten sus juicios. Como sugieren estudios recientes, incluso en contextos especializados, la identificación de la autoría resulta problemática cuando se amplía el análisis hacia imágenes de carácter artístico (Cenerini et al., 2025). Así, el desafío no es meramente tecnológico, sino epistemológico y formativo: qué sucede con el juicio visual cuando los indicadores tradicionales de autoría colapsan y cómo se reconfiguran la confianza y la valoración estética en un escenario donde la producción visual ya no es propiedad exclusiva del agente humano.

La investigación se sitúa en la intersección entre cultura visual, estudios sobre creatividad computacional y educación artística, atendiendo a un punto preciso: cómo se organiza el juicio cuando la atribución de autoría se vuelve un dato incierto y, aun así, sigue operando como condicionante interpretativo. La evidencia experimental en estética y cognición visual muestra que la información contextual sobre procedencia o estatus puede alterar la evaluación incluso cuando el estímulo permanece idéntico, lo que indica que la valoración no se deriva únicamente de propiedades visibles, sino también de marcos semánticos y expectativas sobre origen (Kirk et al., 2009). En el campo específico de imágenes y autoría algorítmica, se ha observado además que el conocimiento o creencia sobre atribución a IAGen modula el juicio sobre la obra, afectando dimensiones como la calidad percibida o la legitimidad estética (Gangadharbatla, 2022). Desde esta perspectiva, el presente estudio aporta evidencias empíricas a un debate sobre la atribución de autoría en imágenes cuya producción no siempre es inferible a partir de sus rasgos visuales. En este marco, el estudio adapta un test visual de Turing para examinar cómo se construye la atribución de autoría bajo incertidumbre perceptiva y cómo el conocimiento contextual puede reordenar la valoración estética. Esta aproximación resulta especialmente pertinente en un contexto de crisis de la evidencia icónica, en el que la imagen ha dejado de funcionar como garante automático de verdad, autenticidad o experiencia vivida. Tal como se ha

señalado desde el ámbito de la teoría de la imagen, la confianza en lo visual se encuentra hoy mediada por dispositivos técnicos, marcos culturales y expectativas aprendidas, más que por la relación directa con un referente verificable (Ritchin, 2013). En este sentido, la aportación del estudio no reside únicamente en medir tasas de acierto, sino en poner de relieve los desplazamientos en los criterios de juicio que acompañan a la expansión de la IAGen en la actual producción visual.

La experiencia estética puede entenderse como un episodio complejo en el que el observador articula procesos de percepción, interpretación y evaluación a partir de un objeto visual determinado. Este episodio no se reduce a la detección de rasgos formales, sino que integra expectativas, conocimientos previos y marcos de atribución que orientan el juicio estético. Desde esta perspectiva, la identificación de una imagen como producto humano o generado por IAGen constituye una operación cognitiva situada dentro de dicho episodio, y no un reconocimiento técnico aislado (Leder y Nadal, 2014).

La noción de confianza se introduce aquí como una dimensión metacognitiva del juicio. No remite al acierto objetivo de la decisión, sino al grado de seguridad que el observador atribuye a su propia identificación. Esta confianza opera como una evaluación de segundo orden sobre el propio proceso de decisión y permite analizar hasta qué punto el sujeto percibe su juicio como sólido, dudoso o revisable. La investigación en metacognición ha señalado que estos juicios de confianza no siempre se alinean con la precisión real, aunque resultan relevantes para comprender cómo se regulan las decisiones en contextos de incertidumbre (Dunlosky y Metcalfe, 2009; Fleming y Lau, 2014; Koriat, 2012). La valoración estética, por su parte, se concibe como una estimación de la calidad artística de la imagen que no depende exclusivamente de sus propiedades visuales intrínsecas. En este sentido, se ha observado que la atribución de agencia —esto es, si se presupone un creador humano o un sistema artificial— puede modular la intensidad de la respuesta emocional, el grado de implicación y la valoración otorgada a la obra. La experiencia estética no se agota, por tanto, en la percepción de forma y contenido, sino que se ve atravesada por supuestos sobre intención, expresividad y origen de la acción creativa (Demmer et al., 2023).

A partir de estas consideraciones, identificación, confianza y valoración se entienden como dimensiones analíticamente distinguibles pero interrelacionadas. La identificación activa inferencias sobre el origen de la imagen; la confianza informa sobre la solidez subjetiva de esa inferencia; y la valoración expresa cómo dicha atribución incide en el juicio estético. Esta articulación permite analizar de manera integrada cómo se construye el juicio visual en un contexto en el que los indicadores tradicionales de autoría han perdido estabilidad y donde la agencia creativa ya no puede darse por supuesta.

2. Marco teórico

2.1. Percepción y detección de imágenes sintéticas

La identificación de la autoría —ya sea humana o generada por inteligencia artificial— descansa sobre inferencias perceptivas que no garantizan una discriminación fiable, incluso ante tareas de elección binaria. En estudios comparativos se observa que la atribución de origen tiende a aproximarse al azar, mientras que la valoración suele mantenerse independiente de la autoría cuando esta no se explicita (Ragot et al., 2020). La evidencia en la detección de imágenes manipuladas apunta a limitaciones persistentes en el rendimiento humano y a una elevada variabilidad interindividual (Bray et al., 2023). Esta dificultad no se explica solo por la complejidad técnica del objeto, sino por la insuficiencia de los criterios intuitivos cuando el material generado satisface expectativas visuales plausibles.

La investigación sobre rostros sintéticos muestra además que la verosimilitud puede alcanzar un umbral donde la lectura humana carece de indicios estables, desplazando el juicio hacia conjeturas basadas en señales débiles o preferencias de respuesta (Lago et al., 2021). Aunque el rostro es un caso sensible por su centralidad perceptiva, estos trabajos describen un patrón general de indistinción aplicable cuando la imagen alcanza una coherencia global suficiente. Asimismo, la detección depende críticamente del tipo de imágenes y de los regímenes de generación evaluados. Al introducir variaciones en los modelos o condiciones de producción, la fiabilidad se reduce y la tarea se vuelve dependiente de distribuciones específicas más que de pistas universales (Lu et al., 2023). Esto resulta clave en dominios

artísticos donde la variabilidad estilística es constitutiva y la ausencia de un patrón único de "imagen humana" debilita los criterios de discriminación.

El análisis forense ha examinado por qué la identificación basada en artefactos visibles pierde eficacia conforme los modelos se perfeccionan (Tolosana et al., 2020). La coherencia en iluminación, textura y estructura espacial reduce la disponibilidad de indicios locales que antes funcionaban como señales de alerta. Bajo este escenario, el juicio depende menos de la detección de errores y más de inferencias sobre el proceso de creación, las cuales pueden ser imitadas por la generación algorítmica. A ello se suman respuestas subjetivas ante lo sintético —como el sublime algorítmico o el siniestro sintético— que median la interpretación de la imagen y su estatuto de agencia (Torres, 2025). Esta inestabilidad se manifiesta también en la detección automática, donde los métodos generalizan mal ante fuentes desconocidas o transformaciones posteriores (Hulzebosch et al., 2020). En respuesta, se han propuesto pruebas inspiradas en el test de Turing que trasladan el énfasis desde las métricas internas hacia el rendimiento en tareas de interpretación bajo incertidumbre (Geman et al., 2015). Estos antecedentes permiten tratar la identificación como un problema de juicio y no como una mera verificación técnica. Incluso en contextos profesionales especializados la discriminación se mantiene limitada frente a imágenes de alta calidad (Shin et al., 2023), lo que sugiere que la experticia no garantiza una atribución fiable cuando el estímulo satisface las expectativas visuales del campo.

2.2. Efecto de la atribución en la valoración artística

La valoración estética no depende solo de cualidades formales, sino también de la atribución de autoría activada al juzgar. Si el observador presupone una creación humana, tiende a interpretar la obra bajo supuestos de intención y esfuerzo; si asume un origen en la IAGen, la lectura se desplaza hacia claves de automatismo y recombinación (Bellaiche et al., 2023).

La atribución funciona como una condición que modifica sistemáticamente la apreciación (Horton Jr. et al., 2023). Presentar una misma obra con etiquetas distintas puede inducir una devaluación de la experiencia estética cuando se asigna a la IA, incluso si los estímulos son idénticos (Chiarella et al., 2022). Este fenómeno revela sesgos antropocéntricos donde la obra se valora por la suposición de una huella humana. Así, la atribución de autoría opera como un marco normativo que organiza la lectura antes de atender a los detalles formales (Millet et al., 2023). En términos analíticos, podemos distinguir dos niveles: el plano perceptivo (composición, técnica, estilo) y el plano atributivo (agencia presupuesta). Este segundo plano tiene la capacidad de reconfigurar el primero, determinando qué se considera expresivo o valioso sin que el estímulo cambie.

2.3. Indicadores y heurísticos de lectura visual

La atribución suele apoyarse en indicios interpretados como señales de plausibilidad. Históricamente, ciertos rastros vinculados a la edición o generación han servido como alertas perceptivas (Korus, 2017), pero esta cartografía es inestable. Conforme los artefactos diagnósticos desaparecen, el juicio se desplaza hacia patrones sutiles y dependientes del contexto (Akhtar et al., 2023). En el ámbito de la evaluación de detectores, este desplazamiento se observa también como un aumento de la dificultad a medida que crece el realismo perceptivo: el índice V_AI propuesto por Imanpour et al. (2025) se relaciona de forma inversa con la precisión de detección, lo que sugiere que las imágenes visualmente más realistas tienden a ser más difíciles de identificar como sintéticas. Por tanto, el problema no se resuelve acumulando rasgos, sino mediante la integración de estos en una interpretación coherente.

En el registro artístico, la lectura se organiza como una reconstrucción de la historia generativa: el observador intenta inferir qué operaciones produjeron la imagen y qué restricciones sugiere su estructura (Tang et al., 2023). Los heurísticos funcionan aquí como reglas de economía interpretativa. Su fuerza no reside en un indicio aislado, sino en la convergencia de razones que explican el proceso de producción. Al aplicar el test de Turing al ámbito visual, se identifican repertorios recurrentes: apelaciones a la lógica interna, a la experiencia estética acumulada y a signos de materialidad o gesto (Wang, 2024).

Lo decisivo es la ambivalencia estructural de estos heurísticos: una irregularidad puede leerse como gesto humano o como fallo algorítmico; la limpieza, como destreza o como homogeneización (Shin et al., 2023). El interés teórico radica, finalmente, en comprender cómo el observador transforma señales

equívocas en razones suficientes para construir umbrales de convicción ante una lectura que carece de criterios estables.

3. Objetivos

El objetivo general consiste en analizar la dificultad para discernir entre autoría humana e IAGen en un corpus diverso de imágenes, describiendo los criterios de juicio que sostienen dicha atribución en observadores expertos del ámbito creativo.

Como objetivos específicos, se busca evaluar el grado de acierto y caracterizar los patrones de error entendiendo la atribución como una decisión cognitiva sujeta a sesgos y variabilidad. Asimismo, se examina la confianza vinculada a cada elección como un indicador metacognitivo esencial para regular el juicio bajo condiciones de incertidumbre. Un tercer propósito es sistematizar y comparar los heurísticos que justifican estas decisiones para identificar qué marcos de lectura se activan al evaluar el origen de la imagen. Además, se analiza el impacto de la autoría sobre la valoración estética mediante el contraste entre la evaluación inicial y la posterior a la revelación del origen, determinando en qué medida esta información reorganiza el juicio de calidad. Finalmente, se pretende delimitar las implicaciones de estos hallazgos para la lectura visual experta en un escenario de producción híbrida donde los indicadores tradicionales de valor han perdido su estabilidad.

4. Metodología

4.1. Diseño del estudio

Se empleó un diseño experimental cuantitativo intra-sujetos con medidas repetidas por estímulo. El dispositivo se planteó como una prueba de indistinguibilidad perceptiva aplicada a la atribución de autoría, donde la tarea consistía en decidir si una imagen era de autoría humana o generada mediante IAGen, explicitando además la seguridad del juicio y los criterios de soporte. Esta lógica se apoya en antecedentes de pruebas de tipo Turing en el dominio visual (Geman et al., 2015). Cada participante evaluó un conjunto fijo de diez imágenes en un ciclo de dos fases: una inicial, para registrar la atribución y valoración previa, y una posterior, donde se registró el cambio en la valoración tras revelar una etiqueta de autoría conforme a un plan prefijado.

4.2. Participantes y muestreo

La muestra fue intencional y se restringió a perfiles expertos vinculados a profesiones creativas y ámbitos relacionados con la imagen (tabla 1). El reclutamiento se realizó mediante redes académicas y profesionales, complementado con un procedimiento de bola de nieve, bajo el criterio de inclusión de ser profesional en las artes visuales.

Tabla 1. Perfil de la muestra por ámbito principal (N = 28)

Ámbito principal	n (%)	Edad media (años)	Experiencia profesional media (años)	Uso de IAGen índice medio (0–4)	Uso de IAGen moda
Educación artística	11 (39,3)	47,5	16,5	1,09	Esporádico
Historia del arte	8 (28,6)	52,6	17,2	1,25	Nunca
Audiovisual	4 (14,3)	31,0	7,5	0,75	Esporádico
Animación	1 (3,6)	58,0	25,0	1,00	Esporádico
Arte digital	1 (3,6)	43,0	8,0	2,00	Mensual
Dibujo	1 (3,6)	58,0	8,0	0,00	Nunca
Ilustración	1 (3,6)	30,0	15,0	0,00	Nunca
Pintura	1 (3,6)	58,0	15,0	0,00	Nunca

Nota. La edad media aproximada (años) se calculó asignando el punto medio de cada tramo: 18–24 = 21; 25–35 = 30; 36–50 = 43; 51–65 = 58. La experiencia profesional media aproximada (años) se estimó asignando valores representativos a cada tramo: <2 = 1; 2–5 = 3,5; 6–10 = 8; >10 = 15; >20 = 25. El uso de IAGen se resumió mediante (a) un índice ordinal (0–4), codificado como Nunca = 0; Esporádico = 1; Mensual = 2; Semanal = 3; Diario = 4, y (b) la moda, entendida como la categoría más frecuente en cada ámbito. Los porcentajes se calcularon sobre N = 28 y pueden no sumar 100% por redondeo.

Fuente: elaboración propia, 2026

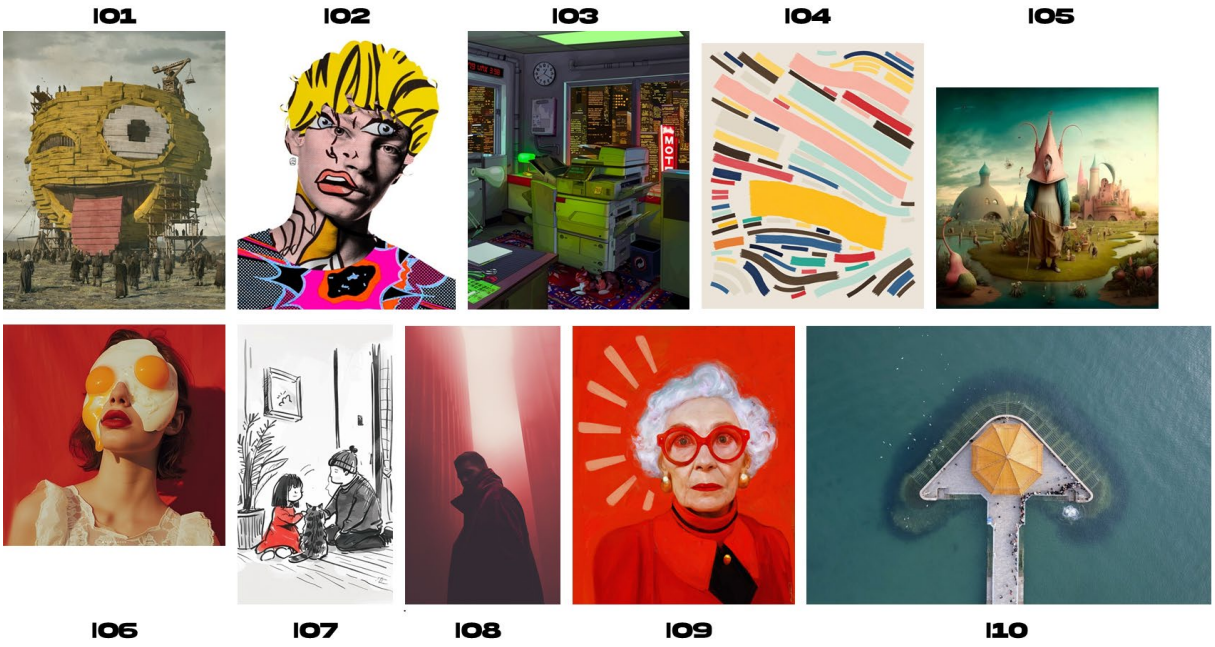
El dispositivo de visionado se registró como covariable contextual: portátil (n = 12), móvil (n = 10) y monitor (n = 6), sin uso de tableta. Las respuestas se recogieron de forma anónima y no se solicitaron identificadores personales.

4.3 Estímulos

El conjunto de estímulos que se muestran en la figura 1, estuvo formado por diez imágenes (I01–I10), equilibradas entre autoría humana (n = 5) y generación mediante IA (n = 5). Para reducir sesgos por diferencias técnicas, las piezas se emparejaron por tema o encuadre, se estandarizaron en resolución y compresión, y se eliminaron sus metadatos. Este procedimiento atiende al riesgo de que la atribución resulte contaminada por rastros no visuales y refuerza la validez del juicio en condiciones de uso real (Hulzebosch et al., 2020).

La selección se diseñó para tensionar la atribución en condiciones de indistinguibilidad perceptiva, integrando lenguajes de la cultura visual contemporánea como fotografía, pintura, ilustración y arte digital. El corpus de IA incluye tanto imágenes generadas mediante Midjourney bajo agencia humana como obras generadas autónomamente a partir del algoritmo "Fidenza", diseñado por Tyler Hobbs. Por un lado, se priorizaron imágenes humanas de factura profesional susceptibles de generar falsos positivos por su acabado digital o carácter extraordinario. Por otro, se seleccionaron piezas de IA de alta calidad para reducir señales algorítmicas evidentes y favorecer falsos negativos. En conjunto, el corpus constituye un set exigente donde la atribución depende de la lectura visual y los heurísticos del participante, y no de marcas externas de procedencia.

Figura 1. Estímulos visuales utilizados en el test



Nota. **I01**: Beeple (2020), *Everydays: The First 5.000 Days* (detalle: emoji); **I02**: Martín Sigwald (2022), *Robert Roy*; **I03**: Mad Dog Jones (2021), *Replicator*; **I04**: Fidenza (2021), *#222*; **I05**: Florelle Guillaume / Midjourney (2023), *El jardín de las delicias* de Jerónimo Bosch en estilo fotográfico; **I06**: @happy_aiccidents / Midjourney (2024), *Sin título*; **I07**: @Tibiami / Midjourney (2024), *Sin título*; **I08**: @Basalte / Midjourney (2024), *Sin título*; **I09**: Alexis Franklin (2024), *Random Portraits II*; **I10**: we-o_rd35hjddgdeuyacd8 (2021), *Sin título*.

Fuente: composición de elaboración propia a partir de las imágenes referenciadas, 2026

4.4 Instrumento

El instrumento se implementó como cuestionario en línea estructurado en: consentimiento, perfil del participante, diez ciclos por imagen y un bloque final de ajuste de criterios. Se configuró con ítems obligatorios y escalas numéricas 1–5. Cada ciclo por imagen se organizó en dos fases. En la fase inicial, se registró la atribución, la confianza (1–5), el gusto, la calidad artística y la pericia percibida. Además, se midió el peso de cinco familias de indicios: estructura y espacio; técnica y materiales; repeticiones y acabados; composición; e iluminación y color. En la fase posterior, tras mostrar la etiqueta de autoría, se midió el cambio en la valoración de calidad mediante una escala ordinal de cinco categorías. El bloque final incluyó un espacio para el motivo principal del acierto y criterios no recogidos previamente.

4.5 Variables

La variable principal fue la identificación de autoría (acierto/error), derivando de ella tasas de falsos positivos y negativos. Se midieron de forma ordinal (1–5) la confianza, la valoración previa y los indicios influyentes. El cambio tras la revelación se registró en una escala de cinco categorías recodificables. Como moderadores se consideraron el ámbito, los años de experiencia y la formación en IAGen. El dispositivo de visionado actuó como covariable contextual.

4.6 Plan analítico

La identificación de autoría se analizará mediante un modelo binomial (regresión logística) con el acierto como variable dependiente. Se introducirán como predictores los pesos de los indicios y los moderadores de perfil, incorporando efectos aleatorios por participante e imagen. La calibración entre confianza y acierto se abordará comparando el rendimiento observado con la confianza declarada. El cambio en la valoración se modelará mediante un modelo ordinal, considerando la interacción entre la autoría mostrada y la real. Se contemplan análisis de robustez por dispositivo y pruebas de sensibilidad ante tiempos de respuesta bajos.

4.7 Control y ética

Se aseguró la comparabilidad técnica de los estímulos y el equilibrio del corpus. La recogida de datos se rigió por los principios de minimización y anonimato, garantizando el derecho de retirada. La gestión de la información se realizó conforme al Reglamento (UE) 2016/679 (GDPR) (European Parliament y Council of the European Union, 2016), con acceso restringido al equipo investigador.

5. Resultados

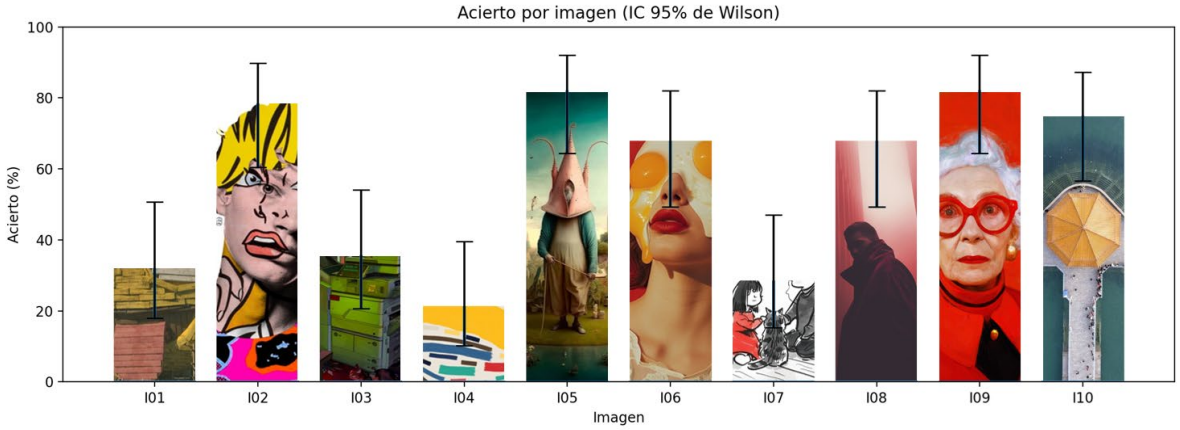
5.1. Identificación de autoría y rendimiento global.

En el conjunto total de decisiones (280 juicios), la tasa de acierto fue del 57,1% (160/280; IC 95% de Wilson [51,3%, 62,8%]). El rendimiento mostró asimetrías por tipo de autoría real. Para las imágenes de autoría humana (5 estímulos; 140 juicios), el acierto fue del 60,7% (85/140). Para las imágenes generadas con IAGen (5 estímulos; 140 juicios), el acierto fue del 53,6% (75/140). En términos de matriz de confusión, se registraron 55 falsos positivos (imágenes humanas clasificadas como IAGen) y 65 falsos negativos (imágenes IAGen clasificadas como humanas). Esto se traduce en una sensibilidad para detectar IAGen del 53,6% (75/140) y una especificidad para reconocer autoría humana del 60,7% (85/140).

5.2. Dificultad por estímulo.

La capacidad de atribución varió de manera marcada entre imágenes. Como muestra la figura 2, las tasas de acierto por estímulo fueron irregulares. En términos de patrón de error, las imágenes humanas I01 e I03 concentraron falsos positivos (19 y 18 clasificaciones como IAGen, respectivamente). Las imágenes IAGen I04 e I07 concentraron falsos negativos (22 y 20 clasificaciones como humanas, respectivamente). Por tanto, el rendimiento global no se explica por una tendencia uniforme, sino por la presencia de estímulos especialmente ambiguos que polarizan el error.

Figura 2. Precisión por nivel de confianza



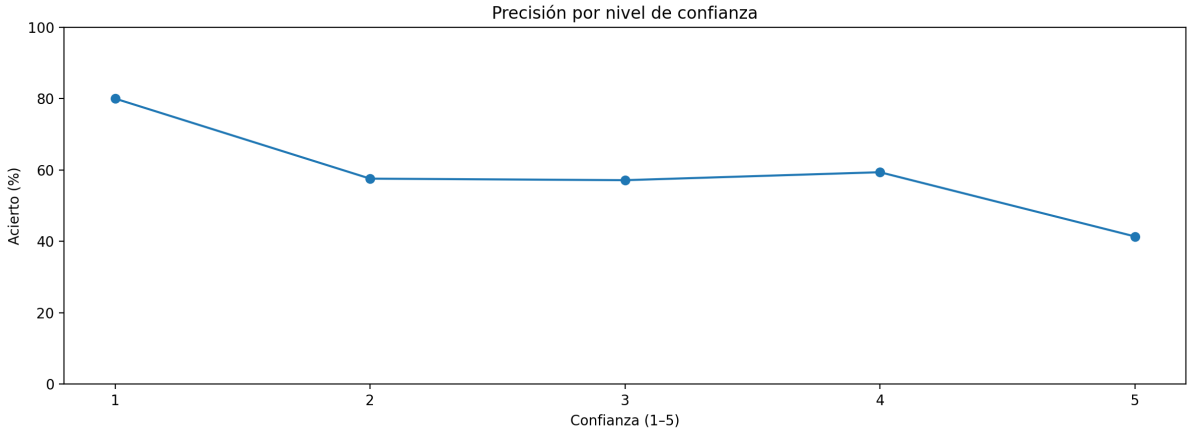
Nota. El eje Y representa el porcentaje de acierto (0–100%) para cada nivel de confianza (1–5). Cada punto corresponde a la proporción de respuestas correctas dentro de ese nivel (n por nivel: 10, 33, 112, 96 y 29, respectivamente).

Fuente: elaboración propia, 2026

5.3. Confianza metacognitiva y calibración.

La confianza media reportada fue 3,36 (DE = 0,94) en una escala de 1 a 5 y no mostró un incremento asociado al acierto. De hecho, la confianza media fue ligeramente mayor en respuestas incorrectas (M = 3,46) que en respuestas correctas (M = 3,29). Al desagregar por niveles, el patrón general no sugiere una calibración monotónica: la precisión se mantiene en torno a valores intermedios entre los niveles 2–4, pero desciende de forma marcada en el nivel máximo de confianza, como se observa en la figura 3. Aunque el nivel 1 debe interpretarse con cautela por su tamaño muestral reducido, el resultado más consistente es la pérdida de precisión en el extremo superior de confianza, compatible con sobreconfianza en una parte de los juicios.

Figura 3. Acierto por imagen (IC del 95% de Wilson)



Nota. El eje Y representa el porcentaje de acierto (0–100%) por estímulo (I01–I10), con 28 juicios por imagen (uno por participante). Las barras de error corresponden al intervalo de confianza del 95% (método de Wilson) para cada proporción de acierto. Las imágenes insertadas en las barras funcionan como referencia visual del estímulo correspondiente.

Fuente: elaboración propia, 2026

5.4. Heurísticos declarados y su relación con el acierto.

Al pedir a los participantes que indicasen, para cada imagen, el peso de cinco familias de indicios en su decisión (escala 1–5), los “motivos predominantes” declarados en los aciertos finales se distribuyeron principalmente entre técnica y materiales (n = 7), iluminación y color (n = 7) y composición (n = 7), con

menor presencia de estructura y espacio (n = 4) y repeticiones o acabados homogéneos (n = 3). Para estimar qué heurísticos se asociaron con una atribución correcta, se ajustó un modelo logístico mixto con acierto (0/1) como variable dependiente, efectos fijos para los cinco pesos y la confianza, e interceptos aleatorios por participante e imagen. Los resultados indican una asociación positiva del acierto con un mayor peso otorgado a estructura y espacio y a iluminación y color, mientras que asignar mayor peso a técnica y materiales, a repeticiones o acabados homogéneos y a composición se asocia con menor probabilidad de acierto; asimismo, una mayor confianza se relaciona con menor acierto (véase Tabla 2).

Tabla 2. Heurísticos y confianza como predictores del acierto

Predictor (peso 1–5)	OR	IC 95% aprox.
Estructura y espacio	1,39	1,29–1,50
Iluminación y color	1,35	1,26–1,45
Técnica y materiales	0,69	0,64–0,74
Repeticiones o acabados homogéneos	0,85	0,79–0,92
Composición	0,82	0,77–0,88
Confianza (1–5)	0,78	0,72–0,84

Nota. Modelo logístico mixto con acierto (0/1) como variable dependiente; efectos fijos para los cinco pesos de indicios y la confianza; interceptos aleatorios por participante e imagen. OR > 1 indica asociación positiva con el acierto; OR < 1, asociación negativa.

Fuente: elaboración propia

La confianza, coherentemente con el análisis descriptivo, mostró una asociación negativa con el acierto en este modelo (OR = 0,78; IC 95% aprox. [0,72, 0,84]). En conjunto, los resultados sugieren que no todos los heurísticos "intuitivamente" plausibles operan como criterios diagnósticos efectivos. Algunos, en particular técnica y materiales, se comportan más bien como una vía sistemática de error dentro de este conjunto de estímulos.

5.5. Valoración de calidad artística y efecto de conocer la autoría.

Antes de conocer la autoría, la valoración media de calidad artística fue 3,11 (DE = 0,90) en escala 1–5. Esta valoración previa fue ligeramente mayor en las imágenes IAGen (M = 3,16) que en las humanas (M = 3,06). Tras revelar la autoría correcta al finalizar la identificación, el cambio declarado en la valoración de la calidad artística se concentró en "no cambia" para ambos tipos de estímulo, pero con diferencias relevantes en la dirección del cambio cuando este se producía (Tabla 3).

Tabla 3. Valoración estética previa y cambio tras revelar la autoría

Tipo de estímulo (autoría real)	Calidad previa (M, DE)	No cambia (%)	Aumenta (%)	Disminuye (%)	Cambio medio (escala 1–5)
Humana	3,06	80,7	10,7	8,6	3,01
IAGen	3,16	73,6	4,3	22,1	2,79
Total	3,11 (0,90)	—	—	—	—

Nota. Calidad previa en escala 1–5. El cambio corresponde a una escala 1–5 (de disminuye mucho a aumenta mucho). Los porcentajes describen la distribución de respuestas de cambio por tipo de estímulo.

Fuente: elaboración propia, 2026

En imágenes humanas, el 80,7% indicó no cambio, el 10,7% indicó aumento (aumenta o aumenta mucho) y el 8,6% indicó disminución (disminuye o disminuye mucho). En imágenes IAGen, el 73,6% indicó no cambio, el 4,3% indicó aumento y el 22,1% indicó disminución. En promedio, la puntuación de cambio fue 3,01 para imágenes humanas y 2,79 para imágenes IAGen (escala 1–5 de disminuye mucho a aumenta mucho), lo que muestra un desplazamiento neto hacia la devaluación cuando el origen se atribuye a IAGen y se conoce como tal.

5.6. Diferencias descriptivas por perfil profesional.

A nivel descriptivo, la precisión media por participante fue 0,57. Por ámbitos principales, las medias fueron 0,57 en educación artística (n = 11), 0,59 en historia del arte (n = 8) y 0,55 en audiovisual (n = 4). El resto de las categorías aparecen con n = 1 y no permiten interpretación. Por uso declarado de

IAGen, las medias fueron 0,58 en “nunca” (n = 10) y 0,57 en “esporádico” (n = 12), con valores algo menores en “mensual” (n = 3) y “semanal” (n = 2). Por años de experiencia, las medias oscilaron entre 0,50 y 0,62 según tramos, sin un gradiente claro. Dado el tamaño muestral y el desbalance entre categorías, estas comparaciones se reportan como descriptivas.

6. Discusión

La atribución de autoría humana o de IAGen, incluso realizada por perfiles expertos vinculados a profesiones de la imagen, no se comporta como una capacidad estable, sino como un juicio situado cuya fiabilidad fluctúa de manera marcada en función del estímulo. Esta variación impide tratar la tasa global de acierto como una propiedad del observador en abstracto y obliga a leer el rendimiento como una relación entre un repertorio de expectativas visuales y un conjunto específico de imágenes. La literatura sobre percepción de imágenes sintéticas ya había mostrado que, cuando la coherencia global del estímulo es suficientemente alta, la discriminación humana pierde apoyos fiables y se desplaza hacia conjeturas sustentadas en señales débiles, incluso cuando la tarea se formula explícitamente como detección (Nightingale y Farid, 2022). En consecuencia, el interés no reside únicamente en constatar que la atribución es difícil, sino en observar cómo se distribuye el error y qué tipos de imágenes activan, en lectores expertos, formas recurrentes de confusión.

La distribución del acierto por imagen ofrece precisamente ese punto de apoyo: ciertos estímulos concentran falsos positivos, es decir, imágenes humanas atribuidas con alta frecuencia a IAGen, mientras que otros concentran falsos negativos, imágenes de IAGen aceptadas como humanas. En línea con el marco teórico, estos desplazamientos sugieren que la atribución tiende a resolverse por plausibilidad y no por detección forense, de modo que el error se concentra allí donde el estímulo activa expectativas estables sobre lo algorítmico o, por el contrario, neutraliza las señales que el observador espera hallar.

En esa misma dirección, los pesos asignados a las familias de indicios permiten ir más allá de una explicación puramente tecnológica y situar el problema en la arquitectura del juicio. Los resultados muestran que no todos los heurísticos que los participantes declaran ponderar se asocian a un mejor rendimiento: algunos se vinculan de forma consistente con una probabilidad mayor de error, mientras que otros se vinculan con una probabilidad mayor de acierto. Este punto es sustantivo porque obliga a desplazar la atención desde el inventario de indicios hacia su rendimiento: importa explicar por qué ciertas señales, aun verosímiles en el imaginario de lo algorítmico, conducen de forma sistemática a atribuciones erróneas. Aquí conviene mantener una cautela conceptual: el instrumento registra autoinformes sobre lo que se cree haber usado como razón, no una trazabilidad directa del proceso perceptivo. Aun así, el patrón es suficientemente robusto para sostener que la lectura experta puede quedar capturada por atajos interpretativos que funcionan como racionalizaciones convincentes, pero no como criterios diagnósticos eficaces. En términos de comparación con el estado de la cuestión, esta inestabilidad de las señales y su dependencia del contexto conecta con síntesis que subrayan que los repertorios de indicios no son universales y cambian con los generadores, las transformaciones y los modos de circulación, de modo que la acumulación de rasgos no equivale a una mejora real del juicio. En registros artísticos, además, la inferencia suele orientarse menos a detectar errores y más a atribuir proceso, puesto que se mira para inferir cómo se hizo y no solo qué hay, lo que favorece explicaciones causales plausibles, aunque no discriminantes.

La dimensión metacognitiva refuerza esta lectura en la medida en que la confianza no actúa como un indicador fiable del acierto y la máxima confianza puede coexistir con errores. Esta descalibración es teóricamente relevante porque muestra que el problema no es únicamente perceptivo, sino también un problema de regulación del propio juicio y de evaluación de la solidez de las razones invocadas, de modo que, en contextos de incertidumbre, la metacognición no consiste en sentirse seguro, sino en calibrar la probabilidad de estar equivocado. Cuando esa calibración falla, el juicio no solo yerra, sino que se vuelve resistente a la revisión, y la literatura metodológica sobre medición metacognitiva permite interpretar este patrón como un desacoplamiento entre precisión y confianza, especialmente esperable cuando los indicios disponibles son ambiguos o cuando el sujeto cree disponer de criterios firmes que, en la práctica, no discriminan. Para educación artística, esta observación tiene un alcance directo, porque la alfabetización visual no puede reducirse a enseñar repertorios de señales, sino que debe incorporar un trabajo explícito sobre la calibración del juicio, la gestión de la duda razonada y la revisión argumentada de los criterios con los que se decide.

La revelación de la autoría al final del cuestionario añade una capa decisiva, porque muestra que la atribución no afecta solo a la identificación, sino también a la valoración estética. El patrón observado es compatible con la idea de que la agencia funciona como operador evaluativo: saber que una imagen es IAGen tiende a desplazar la valoración hacia la devaluación con mayor frecuencia que cuando se sabe que es humana, incluso si la configuración visual permanece idéntica. Este efecto puede entenderse no solo como rechazo de lo algorítmico, sino también como una preferencia positiva por la agencia humana asociada a supuestos de intención, esfuerzo y expresividad, que reordenan el marco de lectura antes de evaluar los rasgos formales. En consecuencia, la discusión sobre IAGen en educación artística no puede limitarse a la indistinguibilidad técnica, porque el problema es también cultural al incidir en cómo se legitima el valor de una imagen y en qué condiciones se considera portadora de experiencia humana. Desde esta perspectiva, el test visual de Turing no funciona solo como medida de rendimiento, sino como un dispositivo que revela el punto en que la atribución reorganiza el sentido y el valor.

El análisis por ámbitos profesionales no autoriza, con la estructura actual de la muestra, afirmaciones sólidas sobre superioridad de unas categorías sobre otras, por razones estadístico-conceptuales: varias categorías tienen tamaños demasiado pequeños para sostener comparaciones, y la experticia no debería reducirse a una etiqueta nominal. Esta limitación, sin embargo, orienta futuras líneas, ya que sugiere operacionalizar la experticia con medidas más finas, por ejemplo, el tipo de práctica visual dominante, la exposición habitual a imágenes generadas, la primacía de tareas de producción frente a tareas de lectura crítica o la formación específica en IAGen. En este marco, la pericia por sí sola no garantiza una atribución fiable cuando el estímulo satisface las expectativas del campo, lo que refuerza la necesidad de precisar qué se entiende por pericia y en qué condiciones opera.

Las limitaciones del diseño deben explicitarse sin dramatizarlas, teniendo en cuenta que el número de estímulos es reducido y su selección condiciona la forma del error, que el muestreo es intencional y no pretende representatividad poblacional, que existe heterogeneidad de condiciones de visionado capaz de introducir variabilidad y que los heurísticos son autoinformes que capturan racionalizaciones plausibles, no necesariamente los procesos perceptivos efectivos. Estas limitaciones no invalidan el estudio, pero sí delimitan el estatuto de las inferencias, porque se justifican afirmaciones sobre patrones de juicio y sobre el modo en que ciertos criterios se asocian con aciertos o errores en este conjunto de estímulos, mientras que no se justifican generalizaciones fuertes sobre la detectabilidad universal de IAGen en cualquier registro visual. Precisamente por ello, la contribución más sólida es metodológica y teórica, en tanto describe cómo se organiza el juicio experto cuando los indicadores tradicionales de autoría pierden estabilidad y la plausibilidad deja de ser un criterio seguro.

Desde aquí, las líneas futuras se derivan con claridad. En primer lugar, conviene ampliar y estratificar el banco de estímulos por registros visuales y por familias estilísticas, incorporando réplicas internas que permitan distinguir la influencia de la autoría de la influencia del registro y aislar qué tipos de configuraciones generan falsos positivos o falsos negativos de manera consistente. En segundo lugar, resulta necesario refinar la medición de la experticia y de la exposición a IAGen, para que el análisis por perfiles no dependa de categorías con tamaño insuficiente. En tercer lugar, sería pertinente separar con más nitidez los componentes perceptivos y metacognitivos del juicio incorporando medidas de proceso, de modo que el estudio de heurísticos no dependa únicamente del relato post hoc, sino que pueda vincularse a dinámicas observables del juicio bajo incertidumbre.

Las imágenes humanas que resultaron especialmente difíciles (I01 e I03) comparten un acabado digital y una coherencia técnica que hoy se asocian con frecuencia a producción algorítmica, pese a su autoría humana. I01 (escena de construcción del gran rostro-objeto) es verosímil en iluminación, textura y profundidad, pero conceptualmente hiperbólica y demasiado resuelta, de modo que la combinación de escala monumental, acumulación de microelementos y apariencia de render puede activar el heurístico según el cual lo espectacular, hiperdetallado y visualmente consistente tiende a atribuirse a IAGen. En I03 (interior nocturno con estética de ilustración 3D), la limpieza de los gradientes, la homogeneidad del color y la organización de aire cinematográfico del encuadre pueden producir un efecto similar, ya que la ausencia de fricción material (rastros del medio, irregularidades del gesto, accidentes del proceso) no se lee como maestría del oficio, sino como señal de generación. Este falso positivo es coherente con un desplazamiento del juicio desde artefactos locales hacia inferencias globales sobre el proceso y hacia estereotipos de estilo asociados a lo algorítmico, que pueden capturar también imágenes humanas de alta factura.

En el polo complementario, los mayores focos de falso negativo (I04 e I07) parecen operar por el mecanismo inverso, porque la IAGen se vuelve difícil de identificar precisamente cuando el registro visual es compatible con prácticas humanas donde la indeterminación, la economía de medios o la irregularidad son no solo aceptables, sino constitutivas. I04 (composición abstracta de bandas y fragmentos cromáticos) presenta marcas que pueden leerse como decisiones compositivas perfectamente plausibles en abstracción contemporánea: distribución rítmica, paleta controlada e imperfecciones susceptibles de interpretarse como huella deliberada. Cuando el estilo no exige verosimilitud figurativa ni consistencia microanatómica, muchos de los indicios habituales pierden fuerza discriminante y la imagen queda protegida por la amplitud interpretativa del género. I07 (dibujo lineal con escena cotidiana) activa un marco de lectura todavía más propicio al error, porque los rasgos que suelen invocarse como humanos en ilustración (línea suelta, simplificación, economía expresiva, cierta torpeza aparente) son hoy fácilmente emulables por IAGen y, además, cualquier anomalía puede reabsorberse como estilo. En estos casos, el heurístico de materialidad y gesto puede funcionar como atajo persuasivo, pero no diagnóstico: se atribuye como humano aquello que encaja con una gramática de lo dibujado, aunque la fuente sea algorítmica. Esto encaja con la noción de que la lectura visual, en registros artísticos, tiende a reconstruir una historia generativa plausible más que a identificar una procedencia de manera forense, con lo cual la plausibilidad estilística basta para sostener atribuciones erróneas.

Las imágenes donde la IAGen fue identificada con mayor facilidad (I05, y en menor medida I06 e I08) sugieren un conjunto distinto de señales, menos ligado a artefactos puntuales que a una configuración global de extrañeza verosímil. I05 (escena fantástica con personaje central y arquitectura imaginaria) reúne elementos que, en la práctica social reciente, han operado como marcador de sospecha: un imaginario onírico solvente en lo atmosférico, pero con una calidad general de ilustración percibida como homogénea, sin zonas de trabajo diferencial, y con detalles secundarios narrativamente indiferentes o formalmente redundantes. El observador puede percibir así una coherencia que no remite a decisiones jerarquizadas, sino a una producción que satura el mundo con la misma intensidad en todos sus planos. En I06 (retrato fotográfico con elemento surreal) y en I08 (figura en penumbra con cromatismo muy marcado), el juicio tiende a apoyarse en la combinación de pulido y rareza: imágenes que parecen fotográficas por iluminación y textura general, pero cuya resolución formal introduce una extrañeza atribuida a generación, especialmente cuando el acabado resulta demasiado uniforme o altamente estilizado. En estos casos, el criterio no es detectar un fallo, sino inferir el tipo de proceso que produce ese equilibrio entre alta verosimilitud y extrañeza controlada.

En el extremo opuesto, las imágenes humanas con alta identificabilidad (I02, I09 e I10) parecen ofrecer apoyos de atribución más próximos a una lectura de proceso. I02 (retrato híbrido con tramas, superposición y dibujo) expone con claridad una lógica de ensamblaje y decisión gráfica donde conviven registros distintos (fotografía, línea, trama, color plano) sin resolverse en una homogeneización global. Esa heterogeneidad controlada funciona como indicio de autoría humana, en la medida en que hace visible una secuencia de operaciones y elecciones que difícilmente remite a una síntesis única. I09 (retrato pictórico con paleta dominante y signos de pincel) refuerza su lectura humana por un tratamiento matérico reconocible, con transiciones y bordes interpretables como producto de oficio y no como suavizado uniforme. I10 (vista aérea del espigón) añade una verosimilitud indexical que, aunque hoy sea técnicamente emulable, sigue operando culturalmente como garantía práctica, por la presencia de microvariaciones del agua, la escala de figuras diminutas, las irregularidades del entorno y una geometría ligada a una infraestructura real. En términos de heurísticos, estas imágenes favorecen una atribución apoyada en estructura y espacio y en material y acabado como reconstrucción plausible del hacer, más que en la búsqueda de artefactos.

Este patrón sugiere una consecuencia central para la interpretación: la dificultad no se distribuye por grado de realismo de manera simple, sino por el grado en que cada imagen permite, o impide, construir una historia de producción con anclajes relativamente estables. Cuando la imagen ofrece señales compatibles con un proceso humano reconocible (heterogeneidad operativa, jerarquía de decisiones, materialidad interpretada como oficio), la atribución tiende a estabilizarse; cuando la imagen se sitúa en registros donde el estilo absorbe la anomalía (abstracción, línea) o cuando presenta un pulido digital que coincide con el estereotipo actual de lo algorítmico, la atribución se vuelve frágil y los errores se agrupan de forma sistemática. En ese marco, la aportación del análisis por estímulos no es inventariar

indicios, sino mostrar cómo ciertos marcos de lectura, perfectamente razonables, pueden ser epistémicamente inseguros en la cultura visual de la IAGen.

7. Conclusiones

Los resultados permiten responder de forma articulada a los objetivos planteados y, a la vez, precisar el alcance de lo que puede concluirse cuando la atribución de autoría se formula como un problema de juicio visual en condiciones de alta ambigüedad. En ese marco, el objetivo general de analizar la dificultad de atribuir autoría humana o de IAGen en un corpus diverso se cumple al constatar que la identificación no se comporta como una capacidad estable, sino como una decisión variable que depende de manera sensible del estímulo. El primer objetivo específico, estimar el grado de acierto y caracterizar el patrón de errores, queda respondido al observar una dispersión marcada por imagen, con estímulos que concentran aciertos y otros que concentran errores sistemáticos, lo que permite describir no solo el rendimiento global, sino la estructura del error. El segundo objetivo, examinar la confianza como indicador metacognitivo y su relación con el acierto, se cumple al constatar que la confianza declarada no se alinea de forma consistente con la precisión del juicio. El tercer objetivo, identificar y sistematizar los heurísticos empleados para justificar las decisiones, se cumple al mostrar que los participantes movilizan criterios reconocibles y verbalizables que no se traducen necesariamente en discriminación efectiva, haciendo visible la distancia entre razón declarada y acierto. El cuarto objetivo, analizar el efecto de la atribución sobre la valoración estética, se responde al comprobar que la explicitación de la autoría puede modificar la valoración, de modo que el juicio estético no se agota en la configuración formal de la imagen, sino que se reordena por la agencia atribuida. El último objetivo, delimitar implicaciones para la lectura visual experta en un contexto de producción híbrida, se cumple al mostrar que la experticia no elimina la ambivalencia, sino que reorganiza la argumentación del juicio ante la falta de indicadores estables.

El trabajo permite describir empíricamente la atribución de autoría como un problema de juicio visual bajo condiciones de indistinguibilidad, más que como un reto técnico de detección. El artículo aporta, en un mismo diseño, tres dimensiones que a menudo se abordan por separado: la precisión en la identificación, la confianza como evaluación metacognitiva de la propia decisión y la valoración estética modulada por la autoría una vez revelada. Al hacerlo con una muestra de perfiles expertos y con un corpus heterogéneo de registros visuales, los resultados sugieren que la atribución opera mediante heurísticos plausibles, pero no necesariamente eficaces, y que esa discrepancia constituye un dato relevante para estudios visuales y educación artística. En este contexto, el juicio experto no desaparece ante la IAGen, pero tampoco se preserva intacto, sino que tiende a reconfigurarse sobre criterios que resultan convincentes en la medida en que encajan con expectativas culturales ya reorganizadas por la presencia de la IAGen en la iconosfera contemporánea.

La recomendación operativa es tratar la atribución de autoría como un objeto de formación y de investigación, y no como una habilidad que se presupone resuelta por la pericia o por la familiaridad con herramientas. En educación artística, esto implica desplazar el foco desde la enseñanza de indicios hacia el trabajo crítico sobre los marcos de lectura, atendiendo a qué criterios se invocan, qué se da por evidente, qué se convierte en prueba sin serlo y de qué modo la seguridad subjetiva puede funcionar como máscara de una inferencia frágil. El problema no reside únicamente en que la IAGen pueda inducir a error, sino en que lo haga con eficacia precisamente porque se apoya en una predisposición ya instalada: una cultura visual habituada a tratar ciertos signos de autenticidad como si fueran pruebas, cuando en realidad operaban como convenciones estabilizadas. Cuando esos signos pasan a ser reproducibles de manera sistemática, no es que perdamos un criterio, sino que se desvela que el criterio siempre tuvo algo de pacto silencioso, un atajo que confundía plausibilidad con verdad. Por ello, un cierre razonable no consiste en reclamar un retorno a garantías imposibles, sino en aceptar que la pregunta decisiva tenderá a desplazarse, de modo que importará menos quién lo hizo y más qué produce la imagen y con qué finalidad. En este desplazamiento, el juicio se orienta hacia la operación visual que la imagen activa, los efectos que organiza, la economía de sentido que propone y la forma de legitimidad que intenta ocupar, porque cuando la autoría se vuelve un dato intercambiable, casi una etiqueta móvil, el núcleo del juicio deja de ser el origen y pasa a ser la función.

8. Agradecimientos

Esta investigación es fruto del proyecto *ARTificial: Integración de la inteligencia artificial generativa como herramienta creativa en las artes visuales* (REDICE24-3625), financiado por el Instituto de Desarrollo Profesional (IDP).

Referencias

Akhtar, Z. (2023). Deepfakes generation and detection: A short survey. *Journal of Imaging, 9*(1), 18. 10.3390/jimaging9010018

Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., Christensen, A. & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research, 8,* 42. 10.1186/s41235-023-00499-6

Bray, S. D., Johnson, S. D. & Kleinberg, B. (2023). Testing human ability to detect "deepfake" images of human faces. *Journal of Cybersecurity, 9*(1), tyad011. 10.1093/cybsec/tyad011

Carceller, A. (2024). The ARTificial Revolution: Challenges for Redefining Art Education in the Paradigm of Generative Artificial Intelligence. *Digital Education Review, 45*, 84-90. 10.1344/der.2024.45.84-90

Cenerini, C., Keller, F., Pennazza, G., Scantonico, M. & Vollero, L. (2025). Artistic Turing test: The challenge of differentiating human and AI-generated art. *Computers in Human Behavior: Artificial Humans, 6.* 10.1016/j.chbah.2025.100209

Chiarella, S. G., Torromino, G., Gagliardi, D. M., Rossi, D., Babiloni, F. & Cartocci, G. (2022). Investigating the negative bias toward artificial intelligence: The effects of prior assignment of AI authorship on the aesthetic appreciation of abstract paintings. *Computers in Human Behavior, 137.* 10.1016/j.chb.2022.107406

Demmer, T. R., Kühnapfel, C., Fingerhut, J. & Pelowski, M. (2023). Does an emotional connection to art really require a human artist? Emotion and intentionality responses to AI- versus human-created art and impact on aesthetic experience. C*omputers in Human Behavior, 148.* 10.1016/j.chb.2023.107875

Dunlosky, J. & Metcalfe, J. (2009). *Metacognition*. Sage.

European Parliament y Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union,* L119, 1–88.

Fleming, S. M. & Lau, H. (2014). How to measure metacognition. Frontiers in Human *Neuroscience, 8,* 443. 10.3389/fnhum.2014.00443

Gangadharbatla, H. (2022). The role of AI attribution knowledge in the evaluation of artwork. *Empirical Studies of the Arts, 40*(1), 55–74. 10.1177/0276237421994697

Geman, D., Geman, S., Hallonquist, N. & Younes, L. (2015). Visual Turing test for computer vision systems. *Proceedings of the National Academy of Sciences, 112*(12), 3618–3623. 10.1073/pnas.1422953112

Horton Jr, C. B., White, M. W. & Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports, 13,* 19001. 10.1038/s41598-023-45202-3

Hulzebosch, N., Ibrahimi, S. & Worring, M. (2020). Detecting CNN-generated facial images in real-world scenarios. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* 642–643. 10.1109/CVPRW50498.2020.00329

Korshunov, P & Marcel, S. (2018). *DeepFakes: A new threat to face recognition? Assessment and detection*. 10.48550/arXiv.1812.08685

Imanpour, N., Borah, A., Bajpai, S., Ghosh, S., Sankepally, S. R., Abdullah, H. M., Kosaraju, N., Dixit, S., Aziz, A., Biswas, S., Jain, V., Chadha, A., Wang, S., Sheth, A. & Das, A. (2025). The Visual Counter Turing Test (VCT^2): A benchmark for evaluating AI-generated image detection and the Visual AI Index (V_AI). En *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics,* 1847–1862. The Asian Federation of Natural Language Processing y Association for Computational Linguistics. https://aclanthology.org/2025.ijcnlp-long.100/

Kirk, U., Skov, M., Hulme, O., Christensen, M. S. & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *NeuroImage, 44*(3), 1125–1132. 10.1016/j.neuroimage.2008.10.009

Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction, 22*(4), 296–298. 10.1016/j.learninstruc.2012.01.002

Korus, P. (2017). Digital image integrity–a survey of protection and verification techniques. *Digital Signal Processing*, *71*, 1-26 10.1016/j.dsp.2017.08.009Get rights and content

Lago, F., Pasquini, C., Böhme, R., Dumont, H., Goffaux, V. & Boato, G. (2021). *More real than real: A study on human visual perception of synthetic faces*. 10.48550/arXiv.2106.07226

Leder, H. & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode—Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4). 10.1111/bjop.12084

Lu, Z., Di Huang, Lei Bai, Jingjing Qu, Chengyue Wu, Xihui Liu & Wanli Ouyang (2023). Seeing is not always believing: Benchmarking human and model perception of AI-generated images. 10.48550/arXiv.2304.13023

Millet, K., Buehler, F., Du, G & Kokkoris, M. D. (2023). Defending humankind: Anthropocentric bias in the appreciation of AI art. *Computers in Human Behavior, 143,* 107707. 10.1016/j.chb.2023.107707

Nightingale, S. J. & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences, 119*(8), e2120481119. 10.1073/pnas.2120481119

Ragot, M., Martin, N. & Cojean, S. (2020). AI-generated vs. human artworks: A perception bias towards artificial intelligence. En *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. 10.1145/3334480.3382892

Ritchin, F. (2013). Bending the frame: Photojournalism, documentary, and the citizen. Aperture.

Shin, K., Lee, J. S., Lee, J. Y., Lee, H., Kim, J., Byeon, J.-S., Jung, H.-Y., Kim, D. H. & Kim, N. (2023). An image Turing test on realistic gastroscopy images generated by using the progressive growing of generative adversarial networks. *Journal of Digital Imaging, 36*, 1760–1769. 10.1007/s10278-023-00803-2

Tang, N., Gong, S., Zhou, J., Shen, M., & Gao, T. (2023). Generative visual common sense: Testing analysis-by-synthesis on Mondrian-style image. *Journal of experimental psychology. General*, *152*(10), 2713–2734. 10.1037/xge0001413

Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A. & Ortega-Garcia, J. (2020). DeepFakes and beyond: A survey of face manipulation and fake detection. I*nformation Fusion, 64,* 131–148. 10.1016/j.inffus.2020.06.014

Torres, A. (2025). A cartography of ambivalence: The sublime and the uncanny in the generative artificial intelligence iconosphere. *Rita*, (24), 65–78.

Verdoliva, L. (2020). Media forensics and deepfakes: An overview. *IEEE Journal of Selected Topics in Signal Processing, 14*(5), 910–932. 10.1109/JSTSP.2020.3002101

Wang, C. (2024). Human or algorithm? The visual Turing test of AI-generated images. *Journal of Multimedia Information System, 11*(3), 201–212. 10.33851/JMIS.2024.11.3.201