

SESIÓN PRINCIPAL · 90 MIN

Instrumentos y análisis de evidencias con IA



Del criterio observable al tablero que orienta una decisión docente

75 minutos de aplicación · 50 registros ficticios · revisión humana obligatoria

El producto final une instrumento, análisis y decisión

A

Instrumento

Criterios observables y anclas inequívocas.

B

Análisis

Denominadores, faltantes y comparaciones prudentes.

C

Decisión

Aceptar, adaptar o rechazar con seguimiento.

LÍMITE DE LA PRÁCTICA

No se cargan datos reales. La IA produce borradores; la lectura de evidencias y la decisión final corresponden a la persona docente.

La sesión dura 90 minutos; 75 son de aplicación

0–5 min	Asistencia y activación	Recuperar el ciclo de evaluación
5–15 min	Microcontenido	Del criterio al dato y a la decisión
15–28 min	Demostración comparativa	Cinco instrumentos, usos y límites
28–50 min	Laboratorio 1	Derivar y revisar un instrumento
50–60 min	Puente hacia los datos	Denominadores, faltantes y calidad
60–78 min	Laboratorio 2	Analizar 50 resultados de tres grupos
78–85 min	Auditoría	Aceptar, adaptar o rechazar la propuesta de IA
85–90 min	Evidencia y cierre	Guardar decisión y nueva evidencia
15 MIN DE ENCUADRE 75 MIN DE APLICACIÓN		

La decisión docente cierra un ciclo y abre el siguiente



Cada instrumento responde una pregunta distinta

Instrumento	Pregunta	Mejor para	Cuidado
Rúbrica analítica	¿Qué tan bien en cada dimensión?	Productos complejos	Adjetivos vagos o falsa precisión
Lista de cotejo	¿Está presente el elemento?	Pasos y requisitos	Confundir presencia con calidad
Escala de valoración	¿En qué grado se observa?	Grado o consistencia	Mezclar atributos en una escala
Guía de observación	¿Qué ocurre durante el desempeño?	Procesos e interacción	Generalizar una observación parcial
Portafolio	¿Cómo progresa a través del tiempo?	Procesos y revisiones	Acumular archivos sin relación

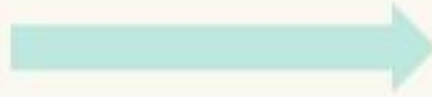
Regla rápida: elige por el tipo de evidencia y la pregunta pedagógica, no por la facilidad de producir un número.

Cuatro niveles, cuatro evidencias, la misma exigencia de trazabilidad

Nivel y caso		Qué observar	Decisión de seguimiento
Educación básica	Historia · cambio y permanencia	Relación entre afirmación y fuente	Modelar el vínculo y volver a redactar
Secundaria	Cívica · sombra o bebederos	Datos, consecuencias y alternativa	Añadir una consecuencia verificable
Media superior	Filosofía · retroalimentación de IA	Tesis, evidencia, objeción y condición	Revisar objeción y condición de uso
Superior	Humanidades · interrupción de estudios	Trazabilidad, límites y pregunta	Delimitar causalidad y buscar otra fuente

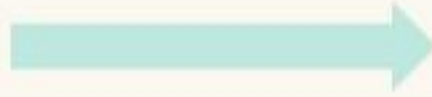
Elige el instrumento por la evidencia, no por comodidad

¿Producto complejo?



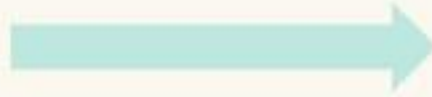
Rúbrica

¿Paso o requisito?



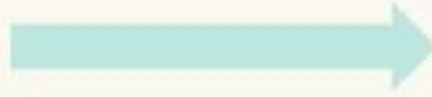
Cotejo

¿Grado observable?



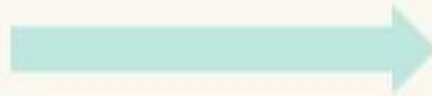
Escala

¿Proceso en vivo?



Observación

¿Progreso en el tiempo?



Portafolio

Después pregunta: ¿qué información perdería si eligiera otra opción?

Las rutas acumulan exigencia: 6, 11 y 20 criterios

BÁSICA

6

criterios acumulados

Instrumento congruente,
denominador, patrón, faltante,
retroalimentación y decisión.

INTERMEDIA

11

criterios acumulados

Suma justificación, anclas claras,
comparación de grupos, acceso y
nueva evidencia.

AVANZADA

20

criterios acumulados

Suma triangulación, regla de
puntuación, distribución, sesgo,
auditoría y reenseñanza.

Elige una sola ruta. Intermedia contiene Básica; Avanzada contiene las dos anteriores.

Laboratorio 1: deriva el instrumento desde la secuencia



Prueba de coherencia: una persona ajena al diseño debe poder localizar evidencia para cada criterio sin adivinar tu intención.

Ejemplo: una rúbrica describe desempeños, no etiqueta personas

MEDIA SUPERIOR · FILOSOFÍA

¿Retroalimentación de IA sin revisión docente?

Anclas 1–4

Inicial · En desarrollo · Logrado · Destacado

K1	Problema y tesis	Delimita el dilema y formula una posición.
K2	Evidencias y razones	Relaciona datos con la tesis y señala límites.
K3	Objeción	Examina una objeción pertinente.
K4	Decisión y revisión	Propone una acción condicionada y revisable.

Describe el mismo atributo en cada nivel y conserva una nota de evidencia localizada.

Un buen prompt conserva límites y revisión humana

1	Contexto	Nivel, propósito, producto y caso ficticio.
2	Criterios	Rasgos observables y anclas solicitadas.
3	Encargo	Instrumento, aplicación, límites y retroalimentación.
4	Prohibiciones	No inventar datos, diagnósticos ni calificaciones.
5	Auditoría	Marcar el resultado como borrador para revisión docente.

Editable · trazable · revisable

SALIDA ESPERADA

Antes de promediar, pregunta qué falta

DATOS FICTICIOS

50

registros totales

48

con evidencia

2

sin evidencia

96 %

participación

Tres comprobaciones antes de analizar

- 1 · ¿Qué significa cada valor?
- 2 · ¿Qué registros faltan y por qué quedan fuera?
- 3 · ¿El denominador acompaña cada porcentaje y promedio?

Los cuatro criterios están muy cerca: K1 merece revisar evidencia

DATOS FICTICIOS · ESCALA 1-4



Qué sí dice

K1 es el menor promedio del conjunto.

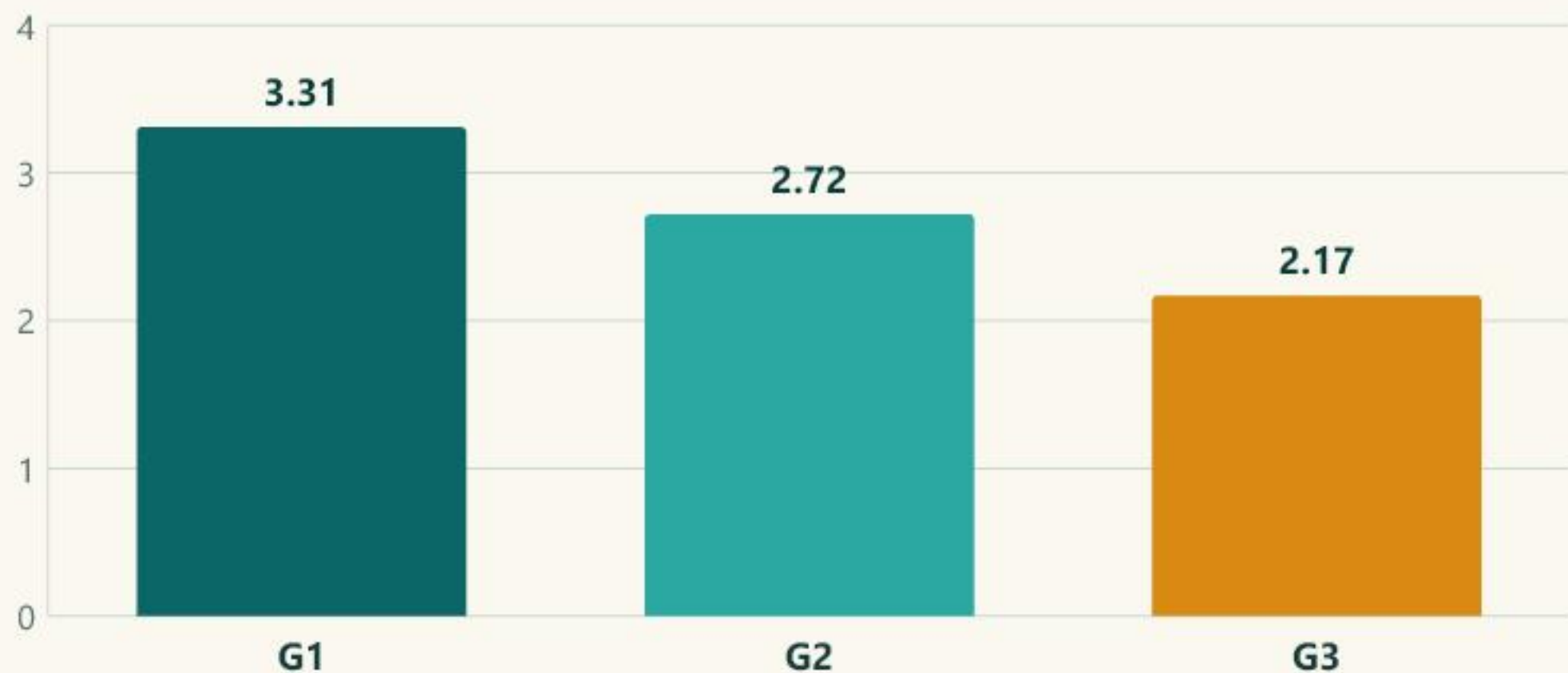
Qué no dice

No explica la causa ni justifica una etiqueta.

La escala inicia en cero para no exagerar diferencias pequeñas.

Los grupos difieren; el tablero no explica por qué

DATOS FICTICIOS · 48 EVIDENCIAS



Comparar no equivale a explicar.

Brecha observada

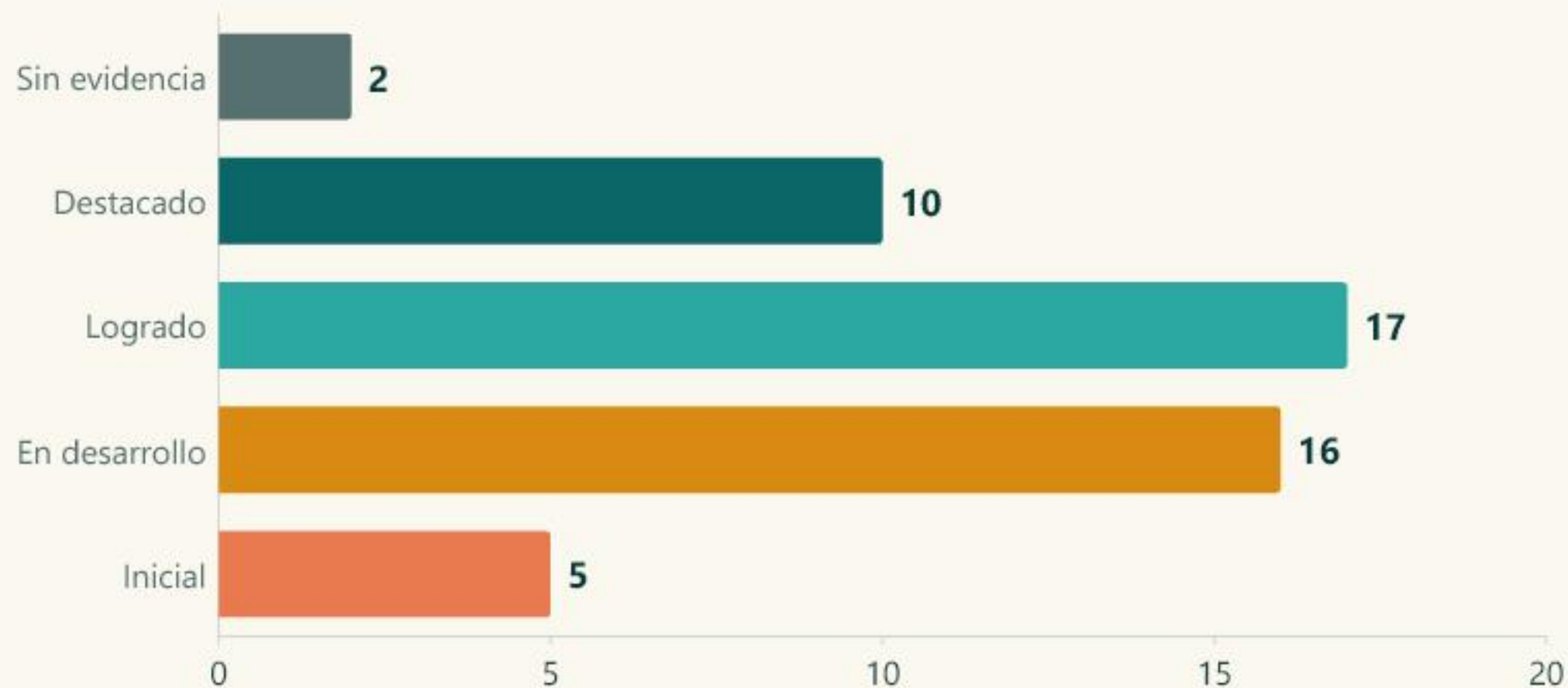
1.14

puntos entre G1 y G3

Pregunta obligatoria: ¿fueron equivalentes las condiciones, el acceso y la interpretación de criterios?

La distribución muestra diversidad y conserva dos ausencias

DATOS FICTICIOS · 50 REGISTROS



Promedio general

2.76 / 4

El promedio no muestra por sí solo la distribución ni las dos ausencias.

Lee siempre promedio + distribución + faltantes.

Separa hallazgo, interpretación y pregunta pendiente

1

Hallazgo

K1 tiene el promedio más bajo: 2.71/4.



2

Interpretación provisional

Podría requerirse apoyo; el promedio no explica la causa.



3

Pregunta pendiente

¿Qué muestran las evidencias originales y las condiciones de aplicación?

Audita el borrador: aceptar, adaptar o rechazar

ACEPTAR

K1 tiene el promedio menor; revisemos las evidencias.

Hallazgo descriptivo y acción prudente.

ADAPTAR

G3 podría requerir apoyo adicional.

Comprobar condiciones y evitar una causa implícita.

RECHAZAR

G3 tiene menor capacidad.

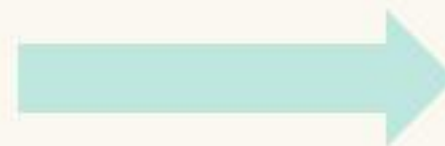
El promedio no autoriza etiquetas ni explica causas.

La revisión humana debe quedar documentada: qué cambió, por qué cambió y qué evidencia se observará después.

La decisión termina en una nueva evidencia

Evidencia acreditable

- ✓ Instrumento y criterios
- ✓ Resumen de 50 datos ficticios
- ✓ Hallazgo y comparación de grupos
- ✓ Decisión docente y limitaciones
- ✓ Adaptación + evidencia de seguimiento



Nueva evidencia

¿Qué observarás después de reenseñar, adaptar o retroalimentar?

Una decisión verificable nombra la acción, su límite y la evidencia que permitirá revisarla.