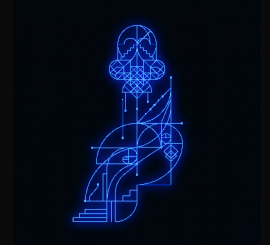


Enseñar a demostrar, no solo a calcular

AP MATH JUSTIFICATION TRAINER

Del argumento empírico —“lo probé con tres números y funciona”— a la justificación formal que un examen AP premia con un 5. El framework CERC vuelve entrenable y medible la argumentación matemática.



ESTADO	STACK	DATOS	MODELO
Beta	Next.js 14 · TS · KaTeX	Mock KV ↔ LMS	scoring CERC

01

EL PROBLEMA

SABER HACER LA CUENTA, NO SABER DEFENDERLA

Los estudiantes de matemática avanzada dominaban el procedimiento pero perdían puntos donde más cuentan: en la justificación escrita. El diagnóstico era nítido y cuantificado. En las plataformas de práctica adaptativa que la red usaba, el rendimiento en opción múltiple rondaba el **82%**; en las Free Response Questions — donde el examinador no evalúa el resultado sino el razonamiento que lo sostiene— caía al **67%**, y solo una minoría iba realmente encaminada hacia la nota objetivo.

La herramienta procedimental existente era sólida para automatizar el cálculo, pero no entrenaba lo que un examen AP de verdad mide: argumentación matemática por escrito.

Un estudiante podía resolver una integral impecablemente y aun así reprobar la pregunta por no enunciar la condición que habilita el teorema que usó.

Ese hueco —saber hacer la cuenta pero no saber defenderla— era invisible para el sistema existente.

Y era, exactamente, la diferencia entre un 3 y un 5.

02

EL OBJETIVO

LA JUSTIFICACIÓN COMO HABILIDAD ENTRENABLE

OBJETIVO PRIMARIO

Llevar al estudiante desde el argumento empírico hasta la justificación formal que cita la hipótesis del teorema y verifica sus condiciones antes de invocarlo — tratando la justificación matemática como una habilidad entrenable y medible, no como un talento difuso.

Para hacerlo operativo se adoptó un andamiaje explícito: el framework **CERC** —Claim, Evidence, Reasoning, Conditions— que descompone toda demostración en cuatro movimientos visibles y exigibles.

La meta no era que el estudiante escribiera más, sino que escribiera *completo*: que ninguno de los cuatro elementos quedara implícito, porque es precisamente lo implícito lo que un examinador no puede premiar.

03

LOS USUARIOS

EL ESTUDIANTE Y EL TUTOR

El usuario primario es el estudiante de matemática avanzada que ya domina el cálculo pero pierde puntos en la respuesta escrita. La plataforma se diseñó para tres cursos paralelos —Calculus AB, Calculus BC y Statistics— porque la estructura del argumento es la misma aunque el contenido cambie: una afirmación, su evidencia, el principio que las conecta y las condiciones que lo habilitan.

El segundo usuario es el tutor o coordinador académico, que opera a través de un panel administrativo con visualización del estado de razonamiento de cada estudiante (R-0388 y sus pares aparecen siempre anonimizados), seguimiento de progreso por unidad y disparadores manuales de práctica.

Una decisión de privacidad deliberada: el panel del tutor aísla los datos sensibles y nunca expone al estudiante las “trampas” pedagógicas de cada problema, para no contaminar el ejercicio.

04

EL ARTEFACTO

LA SESIÓN A PANTALLA DIVIDIDA

La interfaz central es una sesión **split-screen**: a la izquierda el enunciado del problema con la notación renderizada en KaTeX y el recuadro del teorema —nombre, enunciado, hipótesis—; a la derecha el formulario CERC, cuatro campos apilados —Claim, Evidence, Reasoning, Conditions— cada uno con su descripción y su marco de oración cuando corresponde. El contador de completitud y la barra de progreso dan retroalimentación visual en tiempo real a medida que el estudiante completa cada campo.

FIG. 1 – SESIÓN CERC • FRQ DE CALCULUS A LA IZQUIERDA, FORMULARIO CLAIM/EVIDENCE/REASONING/CONDITIONS A LA DERECHA • “3 OF 4 COMPLETE” MUESTRA EL SCORING POR COMPLETITUD EN VIVO

05

EL FEEDBACK

SCORING POR COMPLETITUD, NO POR CORRECCIÓN SEMÁNTICA

Tras el envío, el sistema evalúa de forma **determinista** si los cuatro elementos están presentes y no vacíos, y lo proyecta sobre la rúbrica AP de 1 a 5. No juzga si el argumento es “bueno” con un modelo de lenguaje: el primer hábito que hay que instalar no es la elocuencia, sino la integridad estructural del argumento —que la condición esté escrita, que la evidencia exista, que el razonamiento nombre el teorema.

The screenshot shows a math problem-solving interface. On the left, the problem statement asks to find c in $(0, 2)$ such that $f'(c) = \frac{f(2) - f(0)}{2 - 0}$ for $f(x) = x^3 - 2x + 2$. The right panel, titled 'CERC Framework', provides feedback on the student's solution. It includes sections for Claim, Evidence, Reasoning, and Conditions, each with a score and a comment. The overall score is 73/100, and the student earned +21XP.

Problem Statement
 Consider $f(x) = x^3 - 2x + 2$ on the interval $[0, 2]$. Use the Mean Value Theorem to find all c in $(0, 2)$ such that:

$$f'(c) = \frac{f(2) - f(0)}{2 - 0}$$

State your conclusion using the CERC framework and verify every hypothesis before applying the theorem.

Remember: Check ALL Conditions
 Before applying any theorem, equally verify that all hypotheses are satisfied. This is where most errors occur!

CERC Framework
 Structure your mathematical argument

Claim
 What is your conclusion?
 There exists at least one value c in $(0, 2)$ satisfying the MVT equation; solving gives $c = 2/3 \approx 1.333$.
 Score: 92/100
 Precise, well-quantified conclusion. The exact value of c is correctly identified and stated.

Evidence
 What mathematical data supports your claim?
 $f(2) = 3, f(0) = 2$, so the average rate of change is $(3 - 2)/(2 - 0) = 1/2$. Also $f'(x) = 3x^2 - 2$, so setting $3x^2 - 2 = 1/2$ gives $x^2 = 1/2$.
 Score: 92/100
 Clean computation. Both endpoint values and the derivative are correct and clearly shown.

Reasoning
 Which theorem or principle connects evidence to claim?
 The Mean Value Theorem guarantees a point where the instantaneous rate of change equals the average rate of change over the interval.
 Score: 68/100
 You name the theorem correctly, but the link from "average rate" to your computed c is implied rather than spelled out.

Conditions
 Have you verified ALL theorem hypotheses?
 It is a polynomial, so it's continuous and differentiable. The MVT applies.
 Score: 45/100
 You assert the hypotheses hold but don't verify them explicitly on $[0, 2]$ vs $(0, 2)$. MVT needs continuity on the CLOSED interval and differentiability on the OPEN one — name each separately.

Think about this:
 Your function is a polynomial — but a theorem's hypotheses must be checked at the boundary too. How would you state continuity at the endpoints $x = 0$ and $x = 2$ to make your Conditions step airtight?

OVERALL SCORE
 73/100

XP EARNED
 +21XP

Marks to AP rubric 3/5 — a substantially correct argument with one unclarified condition. Strengthen conditions to meet a 4/5 complete justification.

FIG. 2 – FEEDBACK POST-ENVÍO · COMPLETITUD CERC MAPEADA A LA RÚBRICA AP 1-5 · UN EVALUADOR SEMÁNTICO HABRÍA SIDO MÁS IMPRESIONANTE Y MUCHO MENOS HONESTO SOBRE LO QUE MEDÍA

06

EL FUNDAMENTO

TRES ESTADIOS, CUATRO UNIDADES COGNITIVAS

El diseño se apoya en un modelo de desarrollo del razonamiento matemático con tres estadios sucesivos, codificados literalmente en el tipo de datos de la aplicación: **empírico** (el estudiante se convence con ejemplos), **genérico** (generaliza el patrón pero sin rigor) y **formal** (demuestra invocando la estructura). El curso está hecho para forzar esa transición.

UNIDAD	FOCO COGNITIVO	QUÉ ENTRENA
U1	Romper la ilusión empírica	Problemas diseñados para que la intuición falle: el estudiante experimenta por qué ver tres casos no basta.
U2	Verificación de condiciones	Verificar TODAS las condiciones de un teorema, sin atajos — el error más caro del examen real.
U3	Síntesis sin andamiaje	Síntesis multiconcepto y precisión comunicativa, ya sin marcos de oración.
U4	FRQ cronometradas	Free Response individuales bajo condiciones de examen.

El catálogo clasifica cada problema por el tipo de error que provoca — **CONDITION_BYPASS**, **LOCAL_ONLY_ARGUMENT** o **CER_BREAKDOWN**— y los acompaña de marcos de oración (*sentence frames*) que se van retirando unidad a unidad: una aplicación directa del principio de *fading*, donde el andamiaje se desvanece a medida que la competencia se consolida.

07

EL DISEÑO

EL MODELO CERC Y LA DECISIÓN QUE MÁS IMPORTÓ

CERC descompone toda demostración en cuatro movimientos: la afirmación (Claim), la evidencia que la respalda (Evidence), el principio o teorema que las conecta (Reasoning) y las condiciones que lo habilitan (Conditions). Sistema de pistas de tres niveles —dónde está el fallo, qué elemento CERC está roto, la corrección explícita— sostiene al estudiante sin resolverle el problema. La gamificación suma XP por unidad y desbloquea insignias con animaciones GSAP.

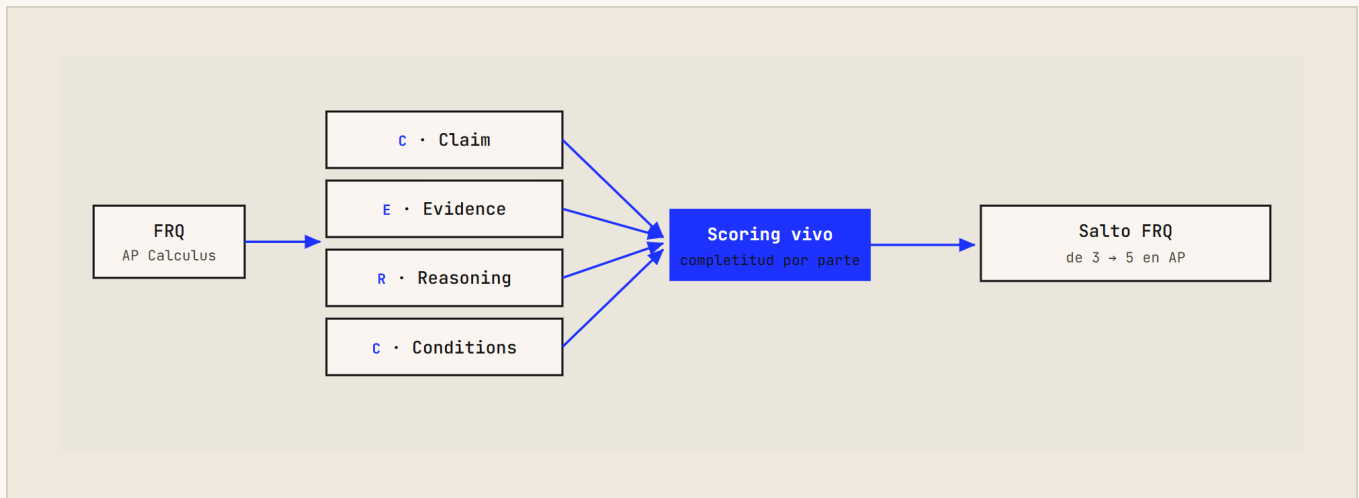


FIG. 3 – MODELO CERC: CLAIM → EVIDENCE → REASONING → CONDITIONS, LOS CUATRO MOVIMIENTOS VISIBLES Y EXIGIBLES DE TODA DEMOSTRACIÓN

Scoring por completitud — no calificación semántica por IA

Contexto. Un evaluador semántico que juzgara si el argumento es “bueno” con un modelo de lenguaje habría sido más impresionante, pero deshonesto sobre lo que realmente medía.

Decisión. El sistema evalúa de forma determinista si los cuatro elementos CERC están presentes y no vacíos, con retroalimentación visual en tiempo real. El primer hábito a instalar es la integridad estructural del argumento, no la elocuencia.

Trade-off aceptado. Mide menos “calidad” y más “completitud” — pero ataca el hábito correcto primero y mide exactamente lo que afirma medir.

La segunda decisión estructural fue la capa de datos con patrón **adaptador**: un adaptador mock para desarrollo respaldado por Vercel KV, y un adaptador del LMS para producción que conserva las formas correctas de los estándares abiertos de interoperabilidad educativa.

08

CONSTRUCCIÓN Y VALIDACIÓN

AI-FIRST, CON LA SEGURIDAD COMO RESTRICCIÓN DE DISEÑO

Se construyó con un flujo **AI-first**, asistido por Claude, a un ritmo deliberadamente alto durante unas seis semanas (17 de marzo al 30 de abril), sobre 29 páginas en App Router de Next.js 14. El stack quedó en

TypeScript estricto, Tailwind, KaTeX para la notación y GSAP para la gamificación.

La seguridad se trató como un frente propio, no como un añadido: autenticación con JWT firmados mediante *jose*, protección de rutas por middleware con control de acceso por rol (estudiante / administrador), sanitización con DOMPurify, rate limiting y CSRF obligatorio en toda mutación autenticada. Una auditoría externa a comienzos de abril detectó oportunidades de endurecimiento en autenticación y manejo de PII; se resolvieron reforzando el modelo de sesión y el aislamiento de datos del panel administrativo.

La validación fue de dos clases. En lo técnico, una batería de pruebas cubría la integridad de los datos del curso, la lógica de prerrequisitos entre unidades y el flujo completo de una sesión, además del build de producción de las 29 páginas. En lo pedagógico, la plataforma se preparó para un **piloto** con un grupo reducido de estudiantes de Calculus BC y Statistics — identificados de forma anónima— con login propio respaldado por el roster real del LMS.

La medición no es una nota subjetiva sino la trayectoria del estudiante a través de los tres estadios de razonamiento y su completitud CERC creciente a medida que el andamiaje se retira entre la Unidad 1 y la Unidad 4.

09

EL RESULTADO

ESTADO Y LEGADO

Una plataforma funcional en estado Beta: las cuatro unidades operativas con su catálogo de problemas, los tres cursos, el motor de scoring por completitud, el sistema de XP e insignias, el panel administrativo con aislamiento de datos y la capa de adaptadores intercambiables. La integración productiva con el LMS quedó como stub deliberado, fiel a las formas los estándares abiertos de interoperabilidad educativa, listo para conectarse cuando el entorno de producción lo permitiera.

El criterio de éxito fijado por el liderazgo académico fue explícito: el trabajo se aprobaba si llevaba al estudiante a un **5** en el examen AP — el outcome, no la actividad.

Framework pedagógico	CERC — Claim · Evidence · Reasoning · Conditions
Unidades · problemas	4 unidades · 7 / 7 / 7 / 4
Cursos cubiertos	Calculus AB · Calculus BC · Statistics
Estadios de razonamiento	empirical · generic · formal (tipo de datos)
Capa de datos	Mock (Vercel KV) ↔ LMS de la red (estándares abiertos de interoperabilidad educativa)
Seguridad	JWT (jose) · middleware por rol · CSRF · DOMPurify · rate limiting
Construcción	AI-first · 29 páginas · capa de adaptadores intercambiables
Estado	Beta · el LMS como stub fiel a las APIs

APRENDIZAJES

- **Honestidad de medición.** Resistir la tentación del evaluador semántico por IA y elegir scoring determinista por completitud: medía exactamente lo que afirmaba medir y atacaba el hábito correcto primero —la estructura antes que la elocuencia.
- **Teoría codificada en los tipos.** Cuando el estadio de razonamiento es un tipo de datos de primera clase y las unidades retiran el andamiaje de forma programada, el diseño pedagógico deja de ser intención y se vuelve verificable.
- **Patrón adaptador.** Decisión de bajo costo y alto retorno: iterar a toda velocidad con un mock persistente sin acoplarse a la API de producción, manteniendo intactas sus formas.
- **Seguridad como restricción de diseño.** En una plataforma con datos de menores, autenticación, CSRF y aislamiento de datos no son una fase final sino una restricción desde la primera línea. la intersección curricular AP/SAT

3

INFRAESTRUCTURA DE CONOCIMIENTO

La capa que fundamenta a todas: de la literatura académica a la investigación con validación de evidencia.

