

No es una opinión sobre cómo se estudia. Es lo que dice la evidencia.

Hacer exámenes en vez de releer, repartirlos en nueve meses, mezclar temas, corregir explicando el error, llevar cuaderno de fallos: cada decisión está tomada de meta-análisis publicados en revistas de referencia.

1.300+

estudios científicos resumidos en los meta-análisis que sostienen el método

110.000+

alumnos participantes en el conjunto de esas investigaciones

8

meta-análisis y revisiones sistemáticas, no estudios sueltos

6

revistas internacionales de primer nivel entre las fuentes, con *Psychological Bulletin* a la cabeza

Tres decisiones sobre cómo se practica

A la izquierda, lo que hace EvaluaMath. A la derecha, el estudio que lo respalda y el tamaño del efecto medido.

1 Hacer exámenes, no releer apuntes

LO QUE HACEMOS

Cada sesión es un **examen completo con el formato de tu comunidad**, generado de nuevo cada vez. El alumno no repasa leyendo: recupera de memoria bajo condiciones de examen.

g = 0,50

de ventaja de practicar con exámenes frente a estudiar leyendo, **medido en aulas reales** sobre 222 estudios y 48.478 alumnos. En una clase de 30, mueve al alumno medio del puesto 15 al puesto 10.

Yang, Luo, Vadillo, Yu y Shanks (2021). «Testing (quizzing) boosts classroom learning». *Psychological Bulletin*, 147(4), 399–435.
UNIVERSITY COLLEGE LONDON · U. AUTÓNOMA DE MADRID

2 Repartidos en nueve meses, no en un atracón final

LO QUE HACEMOS

El plan va de **octubre a junio** y vuelve sobre lo antiguo: un tema no se cierra cuando se aprueba, reaparece semanas después. Es el calendario, no la intensidad, lo que hace el trabajo.

47,3 % frente a 36,7 %

de aciertos con la práctica repartida frente a la concentrada, **con el mismo tiempo total de estudio**. Síntesis de 839 comparaciones en 317 experimentos.

Junto con hacer exámenes, es una de las **dos únicas técnicas de diez** a las que Dunlosky dio **utilidad alta**.

Cepeda, Pashler, Vul, Wixted y Rohrer (2006). *Psychological Bulletin*, 132(3). · **Dunlosky et al. (2013).** *Psychological Science in the Public Interest*, 14(1).
UC SAN DIEGO · KENT STATE · DUKE UNIVERSITY

3 Mezclando temas, no un tema por sesión

LO QUE HACEMOS

Los exámenes de mezcla y los globales **combinan temas distintos**. Así el alumno no practica solo aplicar un método: practica **elegir cuál**, que es exactamente lo que le pedirán en junio.

60,7 % frente a 37,6 %

en el examen final, tras un curso entero de práctica mezclada frente a práctica por bloques. **d = 0,83** · 787 alumnos, 54 clases, institutos públicos.

Honestidad: ese estudio se hizo en **1.º de ESO**, no en bachillerato. En tareas de matemáticas el efecto medio es más modesto, **g = 0,34** (Brunmair y Richter, 2019), pero consistente.

Rohrer, Dedrick, Hartwig y Cheung (2020). *Journal of Educational Psychology*, 112(1), 40–52.
UNIVERSITY OF SOUTH FLORIDA · ENSAYO CONTROLADO ALEATORIZADO

LA PIEZA CON MÁS EVIDENCIA DE TODO EL MÉTODO

4 · Corregir explicando el error vale cuatro veces más que poner la nota

EvaluaMath corrige **apartado a apartado con criterio PAU**: qué has hecho mal, por qué se pierde ese punto y qué tienes que repasar. **Nunca solo la nota**. No es un matiz pedagógico: es el hallazgo más grande de toda esta infografía.

La mayor síntesis publicada sobre retroalimentación ordenó los tipos de corrección por su efecto. La que dice qué, por qué y cómo **cuadruplica** el efecto de limitarse a la calificación.

Wisniewski, Zierer y Hattie (2020). «The power of feedback revisited». *Frontiers in Psychology*, 10:3087. Meta-análisis de 32 meta-análisis · 435 estudios · más de 61.000 alumnos. Universidad de Augsburg · Universidad de Melbourne.

Corrección que explica qué fallaste, por qué y qué hacer

d = 0,99 · 42 estudios

Corrección correctiva: solo si está bien o mal

d = 0,46 · 238 estudios

Solo la nota o un «muy bien / esto está mal»

d = 0,24 · 39 estudios

Y dos consecuencias de hacerlo así

Lo que el método produce como efecto secundario: memoria de los propios errores, y un alumno que llega a junio sin miedo al formato.

5 El cuaderno de errores decide cuál es el siguiente examen

LO QUE HACEMOS

EvaluaMath guarda el **historial de errores de todo el curso** y genera el refuerzo dirigido a ellos. El alumno ve por escrito dónde falla y en qué mejora; el sistema decide qué toca practicar. Eso es **enseñar a regular el propio estudio**, no solo dar más ejercicios.

+8 meses

de progreso académico adicional en un solo curso escolar, en las intervenciones de **metacognición y autorregulación**. Síntesis de **355 estudios**, con calificación de seguridad de la evidencia **alta**. Es de las intervenciones con mejor relación coste-beneficio del catálogo.

Education Endowment Foundation (Reino Unido). «Metacognition and self-regulation», *Teaching and Learning Toolkit*. Organismo público británico de evaluación educativa.
CIFRA REFERIDA AL SISTEMA EDUCATIVO BRITÁNICO

6 Hacer muchos exámenes reduce la ansiedad, no la aumenta

LA OBJECCIÓN DE SIEMPRE

«Ponerle un examen cada semana lo va a poner más nervioso.» Es la duda razonable de cualquier padre. Los datos dicen lo contrario: **lo que da miedo es lo desconocido**, y a la vigésima vez el formato de la PAU ya no lo es.

g = -0,52

de **reducción** de la ansiedad ante los exámenes en los alumnos que practican con pruebas frecuentes de bajo riesgo, frente a los que no. Análisis sobre **24 estudios y 3.374 participantes**.

Condición importante: el efecto aparece con pruebas **frecuentes y de bajo riesgo** —para aprender, no para calificar—. Es exactamente el papel que tiene un examen de EvaluaMath: no cuenta para nada, solo dice dónde estás.

Yang, Luo, Vadillo, Yu y Shanks (2021), análisis sobre ansiedad incluido en el mismo trabajo. *Psychological Bulletin*, 147(4).
UNIVERSITY COLLEGE LONDON · U. AUTÓNOMA DE MADRID

Y LO QUE LA MISMA EVIDENCIA DICE QUE NO FUNCIONA

- ✗ **Releer los apuntes**

Utilidad **baja** en la revisión de Dunlosky. Genera sensación de dominio sin dominio: es la trampa número uno del estudiante de bachillerato.
- ✗ **Subrayar y resaltar**

Utilidad **baja**. Ocupa el tiempo de estudio sin producir recuperación de memoria, que es lo único que consolida.
- ✗ **Corregir poniendo solo la nota**

d = 0,24, el peor de los tres tipos medidos. Un 7,4 sin explicación es información sobre el pasado, no sobre qué hacer.

LO QUE ESTOS ESTUDIOS NO DICEN — Y POR QUÉ LO CONTAMOS

Ninguna de estas investigaciones se ha hecho sobre la PAU ni sobre 2.º de bachillerato español. Lo que está probado son los mecanismos de aprendizaje, no el resultado en esta prueba concreta. EvaluaMath aplica mecanismos validados a un examen que nadie ha estudiado todavía.

En matemáticas los efectos son más modestos que en el conjunto de materias. Un meta-análisis reciente centrado solo en matemáticas encuentra $g = 0,28$ para repartir la práctica y $g = 0,18$, con intervalo que cruza el cero, para la práctica con exámenes (**Murray, Horner y Göbel, 2025, Educational Psychology Review**).

Por eso el método carga el peso donde la evidencia es más fuerte: en la corrección que explica el error ($d = 0,99$), que es la pata más sólida, la mejor medida y la única que no depende de la asignatura. Lo demás suma; eso es lo que decide.

Y TODO ESTO, TRADUCIDO A UNA SEMANA DE TU HIJO

PILARES 1 Y 3

Un examen nuevo, con temas mezclados
Formato exacto de tu comunidad, generado desde cero. **Nunca el mismo.**



PILAR 4

Corrección que explica cada apartado
Qué falla, por qué se pierde el punto y qué repasar. **No solo la nota.**



PILAR 5

El fallo entra en el cuaderno de errores
Queda registrado todo el curso. **El sistema se acuerda en marzo de octubre.**



PILARES 2 Y 6

El siguiente examen va a por ese fallo
Repartido en el tiempo, sin nota que dé miedo. **Y vuelve a empezar.**

Ese bucle **es** el método: la aplicación sistemática, semana a semana durante nueve meses, de cinco cosas que la investigación lleva veinte años midiendo por separado y que casi nadie junta.

REFERENCIAS COMPLETAS

Adesope, O. O., Trevisan, D. A. y Sundararajan, N. (2017). Rethinking the use of tests. *Review of Educational Research*, 87(3), 659–701.
Brunmair, M. y Richter, T. (2019). Similarity matters. *Psychological Bulletin*, 145(11), 1029–1052.
Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T. y Rohrer, D. (2006). Distributed practice in verbal recall tasks. *Psychological Bulletin*, 132(3), 354–380.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J. y Willingham, D. T. (2013). Improving students' learning with effective learning techniques. *Psychological Science in the Public Interest*, 14(1), 4–58.
Education Endowment Foundation. Metacognition and self-regulation. *Teaching and Learning Toolkit*. Londres.
Murray, C., Horner, C. y Göbel, S. M. (2025). Spacing and testing effects in mathematics: a meta-analysis. *Educational Psychology Review*, 37:75.

Rohrer, D., Dedrick, R. F., Hartwig, M. K. y Cheung, C.-N. (2020). A randomized controlled trial of interleaved mathematics practice. *Journal of Educational Psychology*, 112(1), 40–52.
Wisniewski, B., Zierer, K. y Hattie, J. (2020). The power of feedback revisited. *Frontiers in Psychology*, 10:3087.
Yang, C., Luo, L., Vadillo, M. A., Yu, R. y Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning. *Psychological Bulletin*, 147(4), 399–435.