

La recuperación como estrategia pedagógica para la enseñanza de bioestadística

Retrieval practice as a pedagogical strategy for teaching biostatistics

Juan Manuel Solís¹

¹Universidad Nacional de Jujuy, Argentina (juanmasolis@fca.unju.edu.ar)

Cómo citar este artículo:

Solís, J. M. (2024). La recuperación como estrategia pedagógica para la enseñanza de bioestadística. *Educación y ciencia, 13(61), 55-70*.

Recibido: 30 de noviembre de 2023 | Aceptado: 13 de junio de 2024 | Publicado: 15 de julio de 2024

Resumen

El objetivo de este trabajo fue evaluar los efectos de la recuperación como estrategia pedagógica para lograr una mayor apropiación de una serie de conceptos estadísticos seleccionados. La metodología se desarrolló en dos momentos: una primera etapa diagnóstica en la que se comparó los resultados de la recuperación en contraste con el enfoque de enseñanza basada en la repetición de resolución de ejercicios. En una segunda etapa, se implementó la recuperación como estrategia pedagógica de forma sistemática durante la cursada, utilizando algunas de las prácticas sugeridas por la bibliografía consultada (Agarwal et al, 2020; Carpenter y Agarwal, 2020). Los resultados obtenidos en la evaluación en la segunda etapa mostraron una mejora significativa en la apropiación de los conceptos estadísticos.

Palabras clave: recuperación; repetición; estrategia pedagógica; evaluación sorpresa; retención a largo plazo

Abstract

The aim of this study was to evaluate the effects of retrieval practice as a pedagogical strategy to achieve a deeper understanding of a series of selected statistical concepts. The methodology was conducted in two stages: an initial diagnostic phase comparing the outcomes of retrieval practice against a teaching approach focused on repetitive exercise resolution. Subsequently, retrieval practice was systematically implemented throughout the course, incorporating some of the practices suggested by the consulted literature (Agarwal et al., 2020; Carpenter & Agarwal, 2020). The results obtained from the evaluation in the second stage showed a significant improvement in the understanding of the statistical concepts.

Keywords: retrieval; repetition; pedagogical strategy; pop-quiz; long-term retention

INTRODUCCIÓN

La estadística es una disciplina transversal en la investigación científica, y su enseñanza tiene gran relevancia en la formación de profesionales en diversas áreas del conocimiento. En este sentido, resulta fundamental contar con estrategias pedagógicas que permitan a los estudiantes, sin una base consolidada de habilidades analíticas y conocimientos matemáticos previos, apropiarse de conceptos estadísticos de manera efectiva

y a largo plazo, por cuanto en el aprendizaje “las ideas más antiguas poseen ventajas injustas sobre aquellas que llegan más tarde” (Petrosino, 2000).

Tras el confinamiento preventivo por COVID-19, irrumpieron una gran cantidad de recursos y tecnologías de la información y comunicación (TIC) orientados a la enseñanza y el aprendizaje, que abrieron nuevas perspectivas educativas. En particular, herramientas como la librería {exam} de R o los aplicativos interactivos de GeoGebra permitieron ampliar y potenciar los recursos del Aula Virtual en la plataforma Moodle como un complemento versátil y de gran apoyo a las clases presenciales. Siendo así, la cátedra de bioestadística y diseño experimental de la Facultad de Ciencias Agrarias de la Universidad Nacional de Jujuy se dio a la tarea de fortalecer el banco de preguntas del aula virtual a fin de dotar al dictado de la materia un enfoque de enseñanza basado en la repetición de resolución de problemas y ejercicios, de forma tal de propiciar que los alumnos cuenten con un mayor número de intentos en sus prácticas. A pesar de haber concretado un banco de preguntas extenso y variado, lo que permitió que los alumnos pudieran aplicar los conceptos trabajados en clase por medio numerosos ejercicios con diferentes grados de dificultad, los resultados observados durante las evaluaciones no fueron consistentes con la intensidad en el número de intentos, es decir, que la cantidad de prácticas llevadas a cabo no estuvo correlacionada con las calificaciones obtenidas por los alumnos. A raíz de ello, se exploraron alternativas metodológicas para mejorar la capacidad de apropiación de una selección de conceptos estadísticos de los temas considerados de mayor relevancia en: a) Estadística descriptiva; b) Distribuciones de probabilidad; c) Pruebas aplicadas a variables discretas y continuas y d) Inferencia estadística. Entre ellas, el método de la recuperación o retrieval ha sido señalado en diversas fuentes como un dispositivo óptimo para tal fin, ya que fortalece la retención de la información, al tener que recordar y aplicar activamente los conceptos en lugar de simplemente reconocerlos cuando se les presentan.

ANTECEDENTES

Garfield (2008) describe tres niveles de aprendizaje de la Estadística: 1) Alfabetización estadística, b) Razonamiento estadístico y c) Pensamiento estadístico. El nivel básico o alfabetización estadística incluye el entendimiento y el uso básico del lenguaje y las herramientas estadísticas: conocer el significado de términos estadísticos básicos, comprensión del uso de símbolos estadísticos básicos y el reconocimiento e interpretación de diferentes representaciones de los datos. En ese sentido, la alfabetización estadística se relaciona con el nivel de “conocimiento” según la taxonomía de Bloom, para describir el nivel más básico dentro de los resultados del aprendizaje estadístico según su nivel de abstracción (Garfield, 2008), y sobre el cual se apoyan el resto de los niveles o jerarquías de aprendizaje. Desde un enfoque cognitivo, Bransford et al., (2000) señalan que el traspaso analógico exitoso conduce a la inducción de un esquema general para los problemas resueltos que se puede aplicar a problemas subsecuentes. La recuperación de la memoria se promueve mediante esquemas porque estos derivan de un ámbito más amplio de instancias relacionadas que experiencias de aprendizaje únicas, y se potencian mediante el aprendizaje activo, lo que supone un paradigma constructivista del aprendizaje. En este marco, cada nivel de aprendizaje constituye un objetivo a alcanzar habiendo alcanzado competencias suficientes en los niveles previos.

Dentro de los métodos de aprendizaje activos, el método de recuperación o *retrieval* se centra en recuperar información previamente aprendida a partir de la memoria sin pistas ni ayudas externas, fomentando el recuerdo activo siguiendo el principio *no pain, no gain* (Agarwal et al., 2020). A través del acto de recuperación, o “traer” información a la mente, nuestra memoria de esa información se fortalece y es menos probable olvidarla. En contraste, el método de repetición implica revisar y repasar el material de estudio de manera repetida a fin de comprender un concepto determinado o desarrollar una habilidad. Brown, et al. (2014) señalan que “la recuperación repetida puede incrustar tanto el conocimiento como las habilidades de tal manera que se vuelven reflejas” (p.29), siendo más efectiva cuando se divide en períodos separados de entrenamiento. Además, indica que estudiar en forma repetitiva lleva a obtener puntajes más altos en una prueba inmediata, pero resulta en un olvido más rápido en comparación con practicar la recuperación. Esta última afirmación es validada por Marson (2007), quien señala que la repetición mejora tanto la memoria en el corto plazo como la retención y comprensión de un nuevo conocimiento. Por su parte Agarwal (2019, citado en Lang, 2021), indica que la recuperación implica un esfuerzo que es esencial en el proceso de aprendizaje de largo plazo, lo que resulta más efectivo que re-leer, tomar notas o escuchar clases.

Las diferencias en los enfoques de la recuperación y la repetición como estrategias pedagógicas pueden sintetizarse como se indica en la Tabla 1.

Tabla 1

Características distintivas de los enfoques de recuperación y repetición

Característica	Recuperación	Repetición
Descripción	Recuperación activa de la información almacenada en la memoria (“sin ayuda”).	Repetición de ejercicios para reforzar la retención (“con ayuda”).
Retención de conceptos	Largo plazo	Corto plazo
Motivación	Puede mantener el interés al desafiar a los estudiantes y hacer que piensen críticamente	La repetición constante puede llevar a la desmotivación debido a la monotonía
Pensamiento crítico	Desarrolla habilidades de pensamiento crítico ya que la recuperación supone reconocer la aplicación del conocimiento.	Puede no fomentar el pensamiento crítico si se enfoca exclusivamente en la repetición de ejercicios.
Feedback	Se promueve el <i>feedback</i> inmediato y la autoevaluación de procesos.	Feedback y autoevaluación de resultados.
Esfuerzo	Mayor	Menor
Finalidad	Apropiación a largo plazo	Comprensión
Efectividad	Períodos espaciados de entrenamiento	Variabilidad

Nota. Elaboración propia.

La práctica de la recuperación hace al aprendizaje esforzado y desafiante (Agarwal et al, 2020), por lo que su implementación debe ser cuidadosamente planificada y aplicada. En la Tabla 1, se observa que una característica importante a tener en cuenta para lograr una recuperación efectiva es a través de periodos espaciados de entrenamiento (Carpenter y Agarwal, 2020), y por lo tanto su implementación se debe realizar en forma distribuida a lo largo del período educativo a fin de evitar la “ilusión de aprendizaje” que es la percepción de familiaridad de un concepto cuando es tratado en reiteradas ocasiones, cuidando los momentos de aplicación por cuanto es una práctica de alto esfuerzo (Carpenter y Agarwal, 2020).

Se han sugerido diversas estrategias para llevar a cabo una implementación eficaz de la recuperación (Agarwal et al 2022):

- Lluvia de ideas: por medio de tarjetas, pedir a los alumnos que escriban todo lo que recuerden de la clase anterior.
- Cuestionarios sorpresa: Pedir a los alumnos que escriban en menos de un minuto, dos cosas que recuerden luego de una lectura / al final del día / al final de la semana, etc.
- Que los alumnos comenten a sus pares lo que recuerdan sobre un tema.
- Mini-quizzes: preguntas de opciones múltiples o verdadero-falso en tarjetas entregadas a los alumnos a la entrada/salida de clase.

El diseño de entornos educacionales adecuados en los procesos de enseñanza centrada en el estudiante implica identificar y evaluar prácticas pedagógicas que se ajusten a las características y necesidades individuales de los estudiantes. Diversos trabajos de investigación han analizado los beneficios de la recuperación como práctica de aprendizaje (Karpicke, 2012; Karpicke, 2017), pero su aplicación en el aprendizaje de la bioestadística no ha sido suficientemente explorada o, al menos, se ha centrado principalmente en el aprendizaje de la estadística aplicada a ciencias como la psicología (Lyle y Crawford, 2011; Hun Lim, 2015).

Objetivo

El objetivo de este trabajo fue analizar con un enfoque exploratorio los efectos de la recuperación como estrategia pedagógica para lograr una mayor retención a largo plazo, aplicada sobre de una serie de conceptos estadísticos seleccionados en la materia bioestadística de la Facultad de Ciencias Agrarias de la Universidad Nacional de Jujuy: a) Clasificación de variables aleatorias, b) Medidas de variabilidad, c) Independencia de eventos, d) Parámetros de distribuciones de probabilidad, e) Estadístico de prueba en inferencia y, f) P-valor.

METODOLOGÍA

Se realizó un estudio de caso centrado en un análisis exploratorio de los posibles efectos de la recuperación como práctica de aprendizaje aplicado a dos contextos educativos concretos caracterizados por grupos de estudiantes de diferentes carreras. No obstante, la inclusión de un análisis estadístico para evaluar la efectividad de la estrategia de recuperación

intenta añadir mayor objetividad a la investigación, a fin de comparar los resultados obtenidos en cada instancia.

Este trabajo se desarrolló en dos momentos:

1. Momento 1: Una etapa diagnóstica de comparación de la recuperación como estrategia pedagógica para la enseñanza de bioestadística, en contraste con el enfoque de enseñanza basada en la repetición de resolución de problemas y ejercicios, en 27 alumnos de bioestadística, materia del primer cuatrimestre del segundo año de la Licenciatura en Bromatología de la Facultad de Ciencias Agrarias de la Universidad Nacional de Jujuy (Solís, 2023). A este grupo de alumnos se lo denominó Grupo 1.
2. Momento 2: evaluación de los resultados de la implementación de la técnica de recuperación en forma sistemática como estrategia pedagógica en 11 alumnos de bioestadística, materia del segundo cuatrimestre de la Carrera de Licenciatura en Ciencias Biológicas de la Facultad de Ciencias Agrarias de la Universidad Nacional de Jujuy. A este grupo de alumnos se lo denominó Grupo 2.

Es importante notar que la recuperación en el Grupo 1 se realizó como una práctica que consistió únicamente en cuestionarios que los alumnos debieron responder sobre conceptos que fueron recuperados debido a la organización lógica de los temas correlativos en el programa analítico. Estos cuestionarios se basaron en la premisa de que, al recordar y aplicar estos conceptos en diferentes contextos, los estudiantes reforzarían su comprensión y retención de estos. Las preguntas fueron precisas y directas siguiendo las conclusiones de Klahr (2004), quien halló una asociación positiva de la instrucción directa en el aprendizaje. Por ejemplo, la clasificación de variables aleatorias fue un concepto recurrente a lo largo de la cursada, ya que era necesario clasificar las variables antes de aplicar determinadas pruebas estadísticas. Asimismo, el concepto de P-valor fue recuperado en todas las clases donde se realizará una inferencia estadística, a fin de tomar una decisión acerca de rechazar o no las hipótesis nulas.

En cambio, en el Grupo 2 la práctica de recuperación fue diseñada e implementada de forma más estructurada, incorporando algunas estrategias didácticas adaptadas a partir de las provistas en la literatura consultada.

En ambos momentos, utilizamos el software estadístico R para la resolución de situaciones problemáticas.

Momento 1

La cátedra implementó una metodología de aprendizaje basada en la resolución de problemas y ejercicios (Solís, 2023), siguiendo la metodología *Ambiente de Aprendizaje de Razonamiento Estadístico* (SRLE) descrita por Estrella (2014), tratando de integrar el uso de la tecnología con un enfoque activo y participativo de los alumnos poniendo énfasis en la resolución de problemas. Cada clase consistió en la presentación de una situación problemática cuya resolución debía ser propuesta por los alumnos *antes* de desarrollar los temas, según lo observado por Kapur (2014). A continuación, presentar la solución utilizando una metodología estadística adecuada, que luego debía ser aplicada por los alumnos a través

de ejercicios sugeridos de autocalificación en la plataforma Moodle. Para ello aprovechamos la capacidad de ampliar el banco de preguntas de la plataforma Moodle de forma automática con la librería {exams} del programa estadístico R, más específicamente con preguntas condicionales de tipo *cloze*, que son cuestionarios que permiten combinar diferentes tipos de preguntas y respuestas. Por ejemplo, es posible combinar respuestas de tipo numéricas, verdadero/falso y *multiple choice*, entre otras. Además, se puede condicionar una respuesta en función de parámetros previamente especificados, e incorporar plantillas interactivas de GeoGebra para la resolución de ejercicios. {Exams} permite diversificar el Banco de Preguntas en forma aleatoria, creando múltiples cuestionarios a partir de una plantilla. De esta forma, un alumno puede rehacer cuestionarios en múltiples intentos, con resultados diferentes.

A lo largo del curso, realizamos ejercicios de recuperación, es decir, preguntas o ejercicios modelo de clase orientados a que los alumnos traten de recordar una serie seleccionada de conceptos estadísticos con diferentes niveles de dificultad para diferentes temas o unidades curriculares. Por ejemplo, durante la clase práctica de Inferencia con muestras pequeñas, una pregunta fue:

Suponga que el diámetro de colonia de una bacteria coliforme en agar sangre tiene una distribución normal. En un ensayo, la media muestral de los diámetros en 15 colonias fue de 3 mm y un desvío estándar muestral de 0,75 mm. ¿Se trata de datos homogéneos, moderadamente homogéneos o heterogéneos?

En la pregunta anterior, cada alumno debió recuperar el concepto de *Coefficiente de Variación*, introducida durante la clase práctica de Estadística Descriptiva, y aplicarla en el contexto de inferencia de una media, como una forma de anticiparse al comportamiento del intervalo de confianza (a mayor heterogeneidad de datos, se espera una mayor longitud del intervalo de confianza).

Evaluamos cada tema o unidad curricular en dos instancias:

1. Un cuestionario práctico de auto-calificación en la plataforma Moodle, preguntas cerradas, sin restricciones en el número de intentos durante dos semanas luego de finalizada la clase correspondiente a cada tema, computándose la mayor nota obtenida. De esta forma, los alumnos pudieron resolver los ejercicios las veces necesarias hasta lograr la nota de aprobación, o incluso mejorar la nota obtenida. Cada nuevo intento, el juego de preguntas fue diferente por selección aleatoria desde el banco de preguntas.
2. Una instancia de defensa y validación del cuestionario práctico virtual, en forma presencial y oral, donde cada alumno debía desarrollar sobre la resolución de los problemas resueltos.

En la última clase, se les realizó una evaluación sorpresa de 6 preguntas cortas con diferentes niveles de dificultad (Tabla 2):

Tabla 2*Conceptos estadísticos evaluados por tema*

<i>Nº</i>	<i>Concepto estadístico</i>	<i>Tema</i>	<i>Nivel de dificultad</i>	<i>Frecuencia</i>	<i>Puntaje</i>
	<i>Clasificación de v. a.</i>		<i>Muy bajo</i>		<i>1</i>
<i>1</i>	<i>Coefficiente de variación</i>	<i>1</i>	<i>Bajo</i>	<i>6</i>	<i>1</i>
<i>2</i>	<i>Independencia de variables</i>	<i>2</i>	<i>Medio - bajo</i>	<i>1</i>	<i>2</i>
<i>3</i>	<i>Parámetros de la Distr. Binomial</i>	<i>3</i>	<i>Medio</i>	<i>1</i>	<i>2</i>
<i>4</i>	<i>Estadístico Z</i>	<i>5</i>	<i>Medio</i>	<i>1</i>	<i>2</i>
<i>5</i>	<i>P-valor</i>	<i>5 a 9</i>	<i>Alto</i>	<i>5</i>	<i>2</i>

Nota. Elaboración propia.

Los conceptos evaluados tanto en los cuestionarios prácticos como en la evaluación sorpresa correspondieron a los temas Estadística Descriptiva (Tema 1), Probabilidad (Tema 2), Distribución Binomial (Tema 3), Inferencia Estadística con Muestras Grandes (Tema 5) y P-valor (Temas 5 a 9).

La Frecuencia hace referencia al número de clases en los que se recuperó cada Concepto Estadístico, en la forma de un cuestionario práctico de autoevaluación y su correspondiente defensa.

Finalmente, se construyeron las siguientes variables:

- Intentos positivos por tema: la cantidad total de intentos de resolución de cada cuestionario práctico evaluativo, aprobados o no aprobados, realizados por los alumnos. No se incluyen los accesos en los cuales los alumnos no enviaron las respuestas.
- Intentos de superación por tema (intentos promocionados o aprobados): la cantidad de intentos de resolución de cada cuestionario práctico, habiendo obtenido la nota de aprobación con anterioridad.
- Nota promedio de cuestionario práctico por tema: el promedio de calificaciones (en un rango de 1 a 10) obtenido en cada cuestionario práctico.
- Nota promedio de cuestionario práctico sobre intentos aprobados por tema: a diferencia del ítem anterior, no se incluye en el promedio a las calificaciones en intentos desaprobados.
- Nota máxima por cuestionario práctico por tema: la máxima nota obtenida en cada cuestionario práctico por tema.
- Frecuencia por tema: el número de clases donde deliberadamente se realizó al menos una actividad de recuperación de los conceptos estadísticos a ser evaluados.

- Nota de evaluación por tema: el puntaje obtenido en la evaluación sorpresa de cada concepto estadístico.

Las variables “Intentos positivos por tema”, “Intentos de superación por tema”, “Nota promedio de cuestionario práctico por tema” y “Nota promedio de cuestionario práctico sobre intentos aprobados por tema” fueron propuestas como medidas de la motivación de los alumnos hacia la materia. Por ejemplo, se consideró que los alumnos tuvieron un mayor compromiso con el aprendizaje cuanto mayor fue el número de intentos de superación y mayor fue el promedio de cuestionarios aprobados, llegándose a observar alumnos con más de 10 intentos en cuestionarios ya aprobados. Por otro lado, un alto número de intentos con un promedio medio o bajo se asoció a dificultades para aplicar los conceptos tratados en clase.

Los conceptos estadísticos *Clasificación de variables aleatorias* y *Coefficiente de variación* (Tema 1) fueron recuperados a lo largo de 6 clases prácticas a través de preguntas o actividades de clase. El concepto *P-valor* fue recuperado a lo largo de 5 clases prácticas, mediante la resolución de ejercicios de inferencia paramétrica y no paramétrica. El resto de los conceptos estadísticos fueron recuperados únicamente a lo largo de la clase en la que fueron introducidos, pero no se recuperaron en clases subsiguientes.

Se construyó una matriz [27 x 25], con los valores escalados de las variables creadas por tema y alumno/a, excepto las frecuencias de recuperación, sobre la cual se realizó un análisis de componentes principales robustos con el método Hubert mediante la función `pcaHubert()` de la librería `{rrcov}` en R, para evaluar una posible asociación entre la cantidad de repeticiones en la resolución de ejercicios prácticos evaluativos y la calificación obtenida en la evaluación sorpresa.

Se realizó la prueba no paramétrica de Kruskal Wallis sobre la calificación obtenida por tema en la evaluación sorpresa, como respuesta a la frecuencia de recuperación. Para ello, se utilizaron las categorías de Frecuencia de la Tabla 2:

- Nivel 1: se realizaron recuperaciones de los conceptos estadísticos evaluados a lo largo de la clase en la que fueron introducidos, pero no se recuperaron en clases subsiguientes.
- Nivel 5: se realizaron recuperaciones del concepto estadístico evaluado a lo largo de 5 clases.
- Nivel 6: se realizaron recuperaciones del concepto estadístico evaluado a lo largo de 6 clases.

Momento 2

El estudio se realizó sobre el Grupo 2. A partir de los resultados obtenidos durante la evaluación diagnóstica realizada al Grupo 1, en la cátedra decidimos adoptar la técnica de recuperación de forma sistemática como estrategia pedagógica, complementando el enfoque de aprendizaje basado en la resolución de ejercicios. Es decir, cada tema o unidad curricular fue evaluado en dos instancias:

1. Un cuestionario práctico de auto-calificación en la plataforma Moodle, preguntas cerradas, sin restricciones en el número de intentos durante dos semanas luego de finalizada la clase correspondiente a cada tema, computándose la mayor nota obtenida.
2. Una instancia de defensa y validación del cuestionario práctico virtual.

Paralelamente, durante el desarrollo del curso, realizamos las siguientes actividades de recuperación:

- Cuestionarios prácticos de clase: los alumnos debían resolver un cuestionario práctico corto de auto-calificación sobre los temas visto durante en clase las veces necesarias hasta aprobarlo. Este cuestionario podía resolverse en forma individual o en grupos de dos alumnos.
- Cuestionarios prácticos evaluativos acumulativos, es decir, con preguntas vinculadas a temas trabajados en clases previas.
- Ping pong de preguntas no evaluativas: en dos momentos durante el desarrollo del curso, previos a la etapa de defensa presencial de las actividades prácticas, en clase les pedimos a los alumnos que cierren carpetas y apaguen celulares y computadoras. Luego, les proponíamos preguntas sobre todas las unidades vistas hasta ese momento y asignamos al azar a los estudiantes para que desarrollen la respuesta a cada pregunta, realizando una pequeña exposición hacia sus compañeros.
- Preguntas en clase: en cada clase se elegía al azar a uno o más alumnos, y se le/s pedía que explique/n a sus compañeros determinados conceptos estadísticos trabajados durante esa clase, necesarios para resolver los problemas planteados.
- Al inicio de la cursada se realizó un inventario forestal en un macizo boscoso ubicado en la Finca Experimental Emilio Navea, de la Facultad de Ciencias Agrarias de la UNJu, cuyos datos fueron utilizados para la realización de todos los trabajos prácticos de la cursada, lo que permitió vincular los diferentes análisis realizados a un mismo conjunto de datos.

La implementación de la recuperación durante esta segunda instancia fue diferente a la primera respecto de la frecuencia de recuperación. En este caso, el carácter acumulativo de los cuestionarios propició que todos los conceptos estadísticos evaluados sean revisitados en más de una ocasión.

Al finalizar el curso, los alumnos realizaron la misma evaluación sorpresa que el grupo evaluado en primera instancia. En esta instancia, comparamos la calificación obtenida en la evaluación sorpresa entre los alumnos de los Grupos 1 y 2 por medio de una prueba de comparaciones de medias no paramétrica.

Tabla 3

Media aritmética, Mediana (Me), rango intercuartílico (RI) y valor máximo (Máx.) de la nota (en escala 1 a 10) obtenida en la evaluación sorpresa realizada en el Momento 1 por frecuencia de recuperación

Frecuencia	Media	Me	RI	Máx.
1	0,5	0	1	2
5	0,8	0,5	0,75	2
6	0,6	0,5	0,75	1,75

Nota. Elaboración propia.

Tabla 4

Prueba Kruskal Wallis y Wilcoxon – Mann Whitney de la nota obtenida en evaluación sorpresa por frecuencia de recuperación

Prueba	Frecuencias	P - valor
Kruskal Wallis	Todas	0,003 (***)
W – MW	Frec. 1 – Frec. 5	0,012 (**)
(corrección de Holm)	Frec. 1 – Frec. 6	0,017 (**)
	Frec. 5 – Frec. 6	0,35

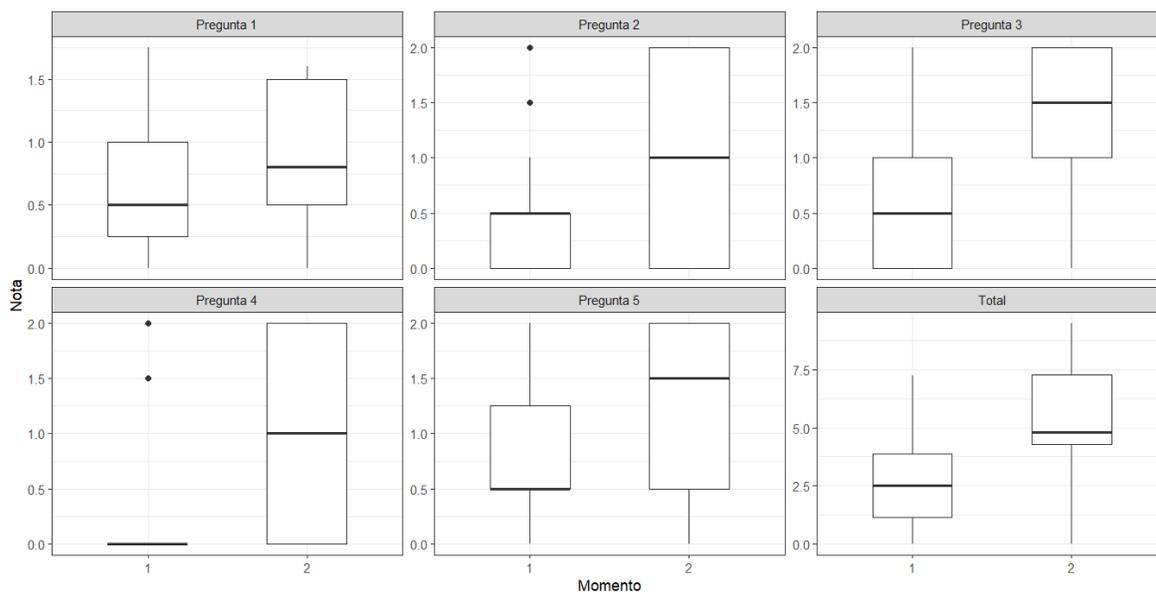
Nota. Elaboración propia.

A partir de la prueba no paramétrica de Kruskal Wallis, observamos diferencias significativas en la nota obtenida en la evaluación sorpresa según la frecuencia de recuperación. La prueba U (W – MW) muestra que existieron diferencias en la nota obtenida en la evaluación sorpresa, cuando se comparó la frecuencia 1 con las frecuencias 5 y 6. No observamos diferencias significativas cuando se compararon las frecuencias 5 y 6 entre sí, a pesar de que el nivel de dificultad de los conceptos estadísticos evaluados en cada caso fue diferente (Tabla 4).

Momento 2

Figura 2

Distribución de la calificación total y por temas obtenida por los grupos evaluados en los momentos 1 y 2



Nota. Elaboración propia.

Tabla 5

Medidas de resumen comparativas de los resultados obtenidos en la evaluación sorpresa por tema y grupo evaluado.

<i>Tema</i>	<i>Grupo</i>	<i>Media</i>	<i>Mediana</i>	<i>RI</i>	<i>Max</i>
<i>Pregunta 1</i>	<i>1</i>	<i>0,63</i>	<i>0,5</i>	<i>0,75</i>	<i>1,75</i>
	<i>2</i>	<i>0,88</i>	<i>0,8</i>	<i>1</i>	<i>1,6</i>
<i>Pregunta 2</i>	<i>1</i>	<i>0,44</i>	<i>0,5</i>	<i>0,5</i>	<i>2</i>
	<i>2</i>	<i>0,95</i>	<i>1</i>	<i>2</i>	<i>2</i>
<i>Pregunta 3</i>	<i>1</i>	<i>0,67</i>	<i>0,5</i>	<i>1</i>	<i>2</i>
	<i>2</i>	<i>1,41</i>	<i>1,5</i>	<i>1</i>	<i>2</i>
<i>Pregunta 4</i>	<i>1</i>	<i>0,28</i>	<i>0</i>	<i>0</i>	<i>2</i>
	<i>2</i>	<i>0,91</i>	<i>1</i>	<i>2</i>	<i>2</i>
<i>Pregunta 5</i>	<i>1</i>	<i>0,8</i>	<i>0,5</i>	<i>0,75</i>	<i>2</i>
	<i>2</i>	<i>1,23</i>	<i>1,5</i>	<i>1,5</i>	<i>2</i>
<i>Total</i>	<i>1</i>	<i>2,81</i>	<i>2,5</i>	<i>2,76</i>	<i>7,25</i>
	<i>2</i>	<i>5,38</i>	<i>4,8</i>	<i>3</i>	<i>9,5</i>

Nota. Elaboración propia.

Las medidas de tendencia central *media aritmética, mediana y valor máximo* de la calificación obtenida fueron superiores en el segundo grupo, tanto a nivel agregado como por temas evaluados. La variabilidad en términos de rango intercuartílico (RI) fue asimismo mayor en la calificación obtenida por el Grupo 2, posiblemente por tratarse de un grupo de menor tamaño.

La diferencia entre las medianas de la calificación fue significativa para los temas Distribución Binomial y Estadístico Z (preguntas 3 y 4), como así también para la calificación total (Tabla 6).

Tabla 6

Prueba de Wilcoxon – Mann-Whitney (U) del contraste de medianas de las calificaciones total y por temas entre grupos.

<i>Tema</i>	<i>W</i>	<i>P-valor</i>
<i>Pregunta 1</i>	<i>107.5</i>	<i>0,1893</i>
<i>Pregunta 2</i>	<i>103.5</i>	<i>0,1285</i>
<i>Pregunta 3</i>	<i>71.5</i>	<i>0,0105</i>
<i>Pregunta 4</i>	<i>91.5</i>	<i>0,0183</i>
<i>Pregunta 5</i>	<i>101.5</i>	<i>0,1263</i>
<i>Total</i>	<i>71.5</i>	<i>0,0138</i>

Nota. Elaboración propia.

Cabe acotar que el resultado agregado o calificación total de la evaluación tuvo una distribución simétrica aproximadamente normal (con un p-valor menor a 0,05 en la prueba de Shapiro Wilk) en ambos grupos, es decir, tuvo un comportamiento robusto y no se vio afectado por valores atípicos.

DISCUSIÓN

En la primera etapa de análisis, la disposición ortogonal en el espacio bidimensional de los grupos A y C, sugiere una falta de asociación entre el número de intentos positivos y el número de intentos de superación con respecto a la nota obtenida en la evaluación sorpresa, por tema (Gráfico 1), lo que indica que la cantidad de intentos realizados no necesariamente se traduce en un aprendizaje efectivo en el largo plazo o comprensión real de los conceptos estadísticos estudiados. Como excepción se puede mencionar la asociación entre el número de intentos positivos y la nota obtenida en la evaluación sorpresa del concepto estadístico del Tema 5, esta última localizada en la región del Grupo A. Asimismo, en términos generales, la calificación obtenida durante los ejercicios prácticos no estuvo asociada a la calificación obtenida durante la evaluación sorpresa. Esta información plantea interrogantes acerca de la suficiencia de las calificaciones parciales obtenidas por los alumnos como método de evaluación del proceso de aprendizaje de los alumnos, e insta a implementar un procedimiento de seguimiento de los alumnos más integral.

Por otra parte, la calificación obtenida en la evaluación sorpresa presentó valores centrales (media y mediana) mayores cuando se evaluaron conceptos tratados con una mayor frecuencia de recuperación. Además, presentaron un menor rango intercuartílico, lo que indica una mayor consistencia en los resultados traducida en una menor variabilidad central de las calificaciones (Tabla 3). La prueba no paramétrica “U” de Wilcoxon - Mann Whitney provee sustento estadístico a la hipótesis de que las calificaciones obtenidas en los conceptos con mayor frecuencia de recuperación fueron mayores a las obtenidas para conceptos no recuperados, pero no fueron estadísticamente diferentes entre sí (Tabla 4). A fin de verificar la validez de estas observaciones, se resolvió continuar con la evaluación del método de recuperación aplicado a un segundo grupo de análisis, lo cual se llevó a cabo en la segunda etapa de análisis.

Si bien las medianas de las calificaciones obtenidas en las preguntas 1, 2 y 5 por el Grupo 2 fueron mayores en comparación con el Grupo 1, estas diferencias no fueron significativas entre ambos grupos (Figura 2 y Tabla 6). Este resultado tiene relevancia dado que las preguntas 1 y 5 corresponden a los temas sobre los que se implementó la práctica de recuperación desde la primera etapa del estudio, lo que puede conducir a la formulación de la hipótesis de que su efectividad podría tener un valor intrínseco más allá de la metodología de enseñanza dentro del cual está incluido. La calificación final mediana fue significativamente mayor en el segundo grupo, lo cual es coherente con los resultados citados en la bibliografía consultada. De esta forma, se pudo corroborar empíricamente que la estrategia de la recuperación en estudiantes de bioestadística permite alcanzar mejoras significativas en la apropiación de conceptos estadísticos en comparación con enfoques tradicionales basados en la repetición de ejercicios, lo cual tiene potencial aplicación en entornos educativos similares.

Además de los resultados observados en las evaluaciones sorpresa, se observó que los estudiantes del Grupo 2 mostraron una mayor atención y participación en clase, como así también compromiso en las actividades de aprendizaje en comparación con el Grupo 1. Durante la validación presencial de las actividades prácticas, los alumnos del Grupo 2 desarrollaron sus exposiciones con mayor seguridad y solvencia para responder. Cabría preguntarse en qué medida estas diferencias pueden atribuirse a la implementación de la práctica de recuperación, y en qué medida a otros factores.

Dado que se trata de un estudio de caso, es conveniente considerar los alcances y limitaciones de las interpretaciones realizadas. En primer lugar, se debe reconocer que los resultados observados en esta instancia no debieran ser generalizables a otros contextos o grupos educativos por cuanto no provienen de un diseño experimental, y se trabajó con diferentes grupos humanos en ambas etapas de análisis. En este sentido, resulta evidente que es necesario continuar con las evaluaciones en la implementación de esta estrategia a fin de aislar variabilidad explicada por factores no controlados que podrían haber afectado los resultados, como la motivación, deseo de autosuperación y situaciones particulares, entre otros. Los efectos de estas variables no controladas pueden estar confundidos en mayor o menor medida con resultados observados y en las diferencias entre los grupos estudiados.

Por lo tanto, los resultados de este estudio deben ser interpretados como hipótesis que requieren ser validadas a través de investigaciones adicionales y repeticiones en otros grupos. No obstante, los resultados han sido coincidentes con las experiencias descritas en otros trabajos de investigación sobre la recuperación, y por tanto constituye una estrategia prometedora que puede ser incluida dentro del diseño de entornos educativos que favorezcan la alfabetización estadística.

CONCLUSIÓN

Los resultados obtenidos en este trabajo sugieren que la práctica de recuperación o *retrieval* es una estrategia efectiva para mejorar el aprendizaje y la retención de conceptos estadísticos, sobre todo si se realiza un diseño estructurado para su implementación en la enseñanza/aprendizaje de bioestadística. El uso de herramientas y recursos TIC como la plataforma Moodle, la librería {exam} de R y los aplicativos de GeoGebra contribuyeron a diversificar el banco de preguntas y construir algunas variables explicativas para analizar los efectos de esta metodología. Estos hallazgos proporcionan una base para futuras investigaciones y para la implementación de prácticas pedagógicas centradas en el recuerdo activo, con el objetivo de fortalecer el aprendizaje a largo plazo. Sin embargo, consideramos necesario continuar con los análisis para determinar si estos resultados son estadísticamente significativos con base en un mayor número de alumnos para aislar un posible efecto grupo y analizar si la estrategia de recuperación influyó en la participación y seguridad de los estudiantes al momento de realizar las actividades.

REFERENCIAS

- Agarwal, P., & Bain, P. (2019). *Powerful Teaching: Unleash the Science of Learning*. Jossey-Bass.
- Agarwal, P., Roediger, H., McDaniel, M., & McDermott, K. (2020). *How to use retrieval practice to improve learning*. Washington University in St. Luis.
- Behar Gutiérrez, R. y Grima Cintas, P. (2004). La estadística en la educación superior ¿Formamos pensamiento estadístico? *Ingeniería y competitividad*, 5(2), 84–90. <https://doi.org/10.25100/iyc.v5i2.2299>
- Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. National Academy Press, Washington, USA.
- Brown, P., Roediger, H., & McDaniel, M. (2014). *Make it stick: the science of successful learning*. The Belknap Press of Harvard University Press.
- Carpenter, S., & Agarwal, P. (2020). *How to use spaced retrieval practice to boost learning*. Iowa State University.

- Estrella, M. S. (2014). Un imperativo moral: La enseñanza de la estadística no puede dejarse al azar. I Encuentro colombiano de educación estocástica. Uniandes Colombia. <https://funes.uniandes.edu.co/funes-documentos/un-imperativo-moral-la-ensenanza-de-la-estadistica-no-puede-dejarse-al-azar/>
- Garfield, J., & Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning: Connecting research and teaching practice*. Springer, Netherlands.
- Kapur, M. (2013). Productive Failure in Learning Math. *Cognitive Science*, 38(5), 1008–1022.
- Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, 21(3), 157–163. <https://doi.org/10.1177/0963721412443552>
- Karpicke, J. (2017). *Retrieval-Based Learning: A Decade of Progress*. Elsevier. Purdue University.
- Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychological Science*, 15(10). <https://doi.org/10.1111/j.0956-7976.2004.00737.x>
- Hun Lim, S., Peng Ng, G. J., & Hao Wong, G. Q. (2015). Learning psychological research and statistical concepts using retrieval-based practice. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01484>
- Lang, J. (2021). *Small Teaching*. 2 Ed. John Wiley & Son.
- Lyle, K. B., & Crawford, N. A. (2011). Retrieving essential material at the end of lectures improves performance on statistics exams. *Teaching of Psychology*, 38(2), 94–97. <https://doi.org/10.1177/0098628311401587>
- Marson, S. M. (2007). Three Empirical Strategies for Teaching Statistics. *Journal of Teaching in Social Work*, 27(3/4), 199-213.
- Petrosino, J. (2000). *¿Cuánto duran los aprendizajes adquiridos?* Ediciones Novedades Educativas. Centro de Publicaciones Educativas y Material Didáctico
- Solís, J. (2023). Avance de investigación sobre la recuperación como estrategia pedagógica para la enseñanza de bioestadística en alumnos la Facultad de Ciencias Agrarias de la Universidad Nacional de Jujuy. *Jornadas Latinoamericanas de investigación en Educación estadística*. Universidad Nacional del Litoral.