



Valoración de una intervención didáctica en medición mediante un diseño pre-experimental

Ignacio Rieiro-Marín

Universidad de Castilla la Mancha, Toledo, España, Ignacio.Rieiro@uclm.es

Melody García-Moya

Universidad de Castilla la Mancha, Toledo, España, melody.garcia@uclm.es

Paloma, Ocaña Aranda

Universidad de Castilla la Mancha, Toledo, España, palomaocara92@hotmail.com

Raquel Fernández-César¹

Universidad de Castilla la Mancha, Toledo, España, Raquel.FCezar@uclm.es

Fecha de recepción: 05-12-2019

Fecha de aceptación: 12-12-2019

Fecha de publicación: 15-12-2019

RESUMEN

En el presente trabajo se analiza la eficiencia de una propuesta de intervención didáctica dirigida a facilitar la superación de dificultades en el aprendizaje de la medida de áreas de superficies y volúmenes de cuerpos para alumnado de Educación Primaria, motivada por diferentes tipos de errores reconocidos. Se empleó un diseño pre-experimental con alumnado de 5º curso. Como instrumentos se utilizaron un pretest y un postest diseñados "ad-hoc". Para objetivar la posible influencia de la intervención, se usó una variable externa, sin relación directa con la intervención. Pese a las limitaciones de los diseños pre-experimentales, se concluye que el presente estudio aporta suficiente información para valorar la utilidad de la intervención didáctica, y postular vías de mejora para orientar un futuro diseño experimental, o adoptar la investigación colaborativa en las aulas usando meta-análisis.

Palabras clave: área, volumen, dificultades de aprendizaje, diseño pre-experimental, Educación Primaria.

Evaluation of a didactic intervention in measurement through a pre-experimental design

ABSTRACT

The present work analyzes the efficiency of a didactic intervention aimed at facilitating the overcoming of difficulties in the learning of the measurement of areas of surfaces and volumes of bodies for Primary Education students, motivated by different types of identified errors. A pre-experimental design is used with 5th grade. Ad-hoc designed pretest and posttest were delivered as instruments. To objectify the possible influence of the intervention, an external variable, which was not directly related to the intervention, was used. Despite the limitations of pre-experimental designs, it is concluded that they provide enough information to assess the usefulness of this intervention, to postulate ways of improvement that could guide a future experimental design, or to adopt collaborative research in the classrooms by using meta-analyses.

Key words: area, volume, learning difficulties, pre-experimental design, Primary Education.

¹ Autora de correspondencia.

1. Introducción

En el presente trabajo se propone un diseño metodológico de investigación en didáctica de las Matemáticas basado en el análisis cuantitativo desde una aproximación a los marcos reales de trabajo en las aulas. Para ello, se afronta el análisis del diseño pre-experimental reforzado con elementos iniciales condicionados por la praxis ordinaria de los centros educativos, con el fin de validar y evidenciar la influencia que la intervención didáctica tiene en la adquisición de elementos competenciales.

La intervención didáctica se diseñó para la superación de algunas dificultades relacionadas con errores tipificados en el aprendizaje de medidas de áreas de superficies y volúmenes de cuerpos con estudiantes de 5º curso de Educación Primaria. Para el análisis se incorporaron elementos relacionados con el tratamiento estadístico de datos cuantitativos y con el diseño de instrumentos, con la finalidad de medir dicha variación con la mínima perturbación en el entorno de un aula educativa.

En el ámbito experimental de la investigación en el aula se percibe cierta dicotomía, como la investigación-acción en el aula contra la investigación de diseño, que se referencia en Romero (2012). Sin embargo, desde este trabajo se asume que existe un único paradigma con dos desarrollos metodológicos diferentes, aunque de similar puesta en valor.

La investigación con análisis de datos cuantitativos en Didáctica de la Matemáticas parece que aboca a un diseño de experimentos. Por tanto, se define dicho término haciendo alusión al trabajo de Montgomery y Douglas (2004), el cual enfoca el diseño de experimentos como una prueba en la que el experimentador controla desde los factores de influencia objeto de estudio hasta los factores no deseados, que, cuando puede, bloquea; o las variables que covarían con los factores. Todo ello se enmarca en un proceso de selección de muestras aleatorias de los sujetos susceptibles de medida, con garantías de control del error entre niveles de los factores, e intraniveles de los factores.

Este marco referencial del diseño experimental deseable se encuentra en contraposición con la realidad de los centros educativos donde la consecución de algunas de estas condiciones garantistas implica una perturbación de la estructura sistémica que hace dudar de que lo que se obtenga sea representativo de la realidad no general, sino solo de la realidad del momento del experimento. Por otra parte, se encuentra la situación que tiene a su alcance el maestro de Educación Primaria (EP) que dispone de capacidad para acceder a la información (Romero, 2012) y así elaborar la situación didáctica de investigación, extrayendo información de la situación ordinaria de desarrollo educativo de forma participativa, colaborativa, crítica, y, por dialéctica, adaptable a las modificaciones de la muestra en la propia interacción de la acción educativa. Esta situación de la investigación-acción en el aula (o en más amplios espacios, como el centro o el municipio), si tuviese una garantía suficiente de reproducibilidad, resultaría más próxima a la realidad que cualquier otra, incluso aunque modificara los límites de la estructura sistémica manteniéndose dentro de ellos, como reconocen Chevallard (1999) y Chevallard et al. (febrero de 2015). La secuencia didáctica debía atravesar transversalmente varios niveles y unidades didácticas, a partir de una selección explícita de situaciones didácticas (Brousseau, 1986). Esta parte de la investigación se concretó en la creación de una "Propuesta de intervención didáctica para superar las dificultades en el aprendizaje de Área y Volumen en Educación Primaria" (Rieiro-Marín, Ocaña, García-Moya, y Fernández-César, 2019).

Dejando la estructura de las ciencias naturales, o de las ciencias exactas, y mirando más a la aplicación y adaptación a la realidad de la Didáctica de la Matemáticas y otras áreas, se hace referencia al libro de Campbell Donald y Stanley (1963) orientado a la educación, psicología, ciencias sociales y ciencias médicas. En este trabajo se percibe un enfoque persistente sobre las amenazas a la inferencia válida, es decir, a la validez de los diseños de pretest y postest de un solo grupo. Se afirma en él que este diseño es de gran aplicación en la investigación educacional. Por otro lado, se encuentra el trabajo de Marsden

y Torgerson (2012), donde critican a algunos autores que utilizaron estos diseños pre-experimentales por no tener en cuenta posibles efectos de sesgo, diciendo que el grupo control aporta aleatoriedad y control de la maduración e historia.

Debido a lo anterior, se encuentran trabajos en Didáctica de las Matemáticas que permiten convertir los datos reportados junto con los parámetros que definen las variables contextuales, en una herramienta básica de colaboración a un posible meta-análisis. Para ello, definen de forma minuciosa el contexto donde se realiza el estudio, las pruebas que se efectúan a los participantes (pretest y postest) y realizan los contrastes pertinentes y analizan los posibles sesgos en el proceso entre ambas pruebas. De esta forma se consigue obtener el control de variables externas sin necesidad de un grupo control. Ejemplos de lo anterior son el estudio realizado por Rojas Álvarez (2013) orientado a medir el efecto de la instrucción geométrica en la representación de módulos multicubos, evaluando y contrastando las respuestas que 17 estudiantes dan al pretest y postest idénticos, y el de Ortiz Aguilar et al. (2017) quienes seleccionan una muestra aleatoria de 30 estudiantes y emplean un pretest, una intervención centrada en el pensamiento geométrico, y un postest que valoraba los efectos de la posible mejora.

En este punto, se pretende saber cuál es la causa de la reiteración en el uso de los diseños pre-experimentales, dado el nulo carácter inferencial de los mismos, y dada la ausencia de aleatorización en general en los mismos. Para poder responder hay que conocer que la educación en España es prácticamente plena de los 3 (96,3%) a los 16 (96%) años. El número medio de alumnos en Educación Primaria en 2019 es de 21,9. En España hay 9.952 centros de enseñanza infantil de primer ciclo, 14.125 de Educación Infantil (EI) de segundo ciclo, y 13.863 de Educación Primaria, la mayoría ofreciendo también Educación Infantil de segundo ciclo, CEIP. De ellos, 10.335 son públicos y 3.538 privados. Los centros públicos de Primaria puros, que son 9.832, tienen 134.562 unidades con 2.787.941 estudiantes, mientras que los centros privados de primaria puros tienen 461 unidades, con 4.774 unidades y con 101.058 estudiantes. Si se añade a estos datos las unidades que tendrían los centros de EP y EI, de una línea (9), de dos líneas (18) y de 3 líneas (27), se puede concretar que la posibilidad de implementar estudios mediante diseños experimentales queda muy limitada al considerar los municipios o centros escolares.

En las tablas 1 y 2, con estimaciones aproximativas, se valora que en los municipios de hasta mil habitantes se tienen grupos de menos de una línea, o sea incompletos. En los municipios de entre 1.000 y 2.000 habitantes se dan grupos incompletos y de una línea, más abundantes los segundos. En los centros de 2.000 a 5.000 se tendrían de línea 2, y de una línea con una proporción de 40% y 60%, respectivamente, donde el 12,73% de la población estaría fundamentalmente en centros de una línea y centros de línea dos. En municipios de cinco mil a diez mil habitantes, habría centros de una, dos o de tres líneas en una proporción estimada de 30%, 60% y 10%, respectivamente. La predominancia de centros de dos líneas es muy alta en el resto, aunque haya muchos centros de 3 líneas a partir de municipios de 25.000 habitantes en adelante.

Habitantes	Todos los centros	E. Infantil	E. Primaria	Primaria/ESO
Hasta 1.000	13,3	4	13,5	18,3
1.001 a 2.000	5	1,7	5,3	10,1
2.001 a 5.000	7	2	8,8	11,8
5.001 a 10.000	10	3,1	12,8	15,6
10.000 a 20.000	13,2	4,3	15,5	17,1
20.000 a 100.000	14,3	4,5	15,8	17,7
100.000 a 500.000	15,4	4,4	16,1	20,1
Más de 500.000	15,6	4,4	15,4	20,8
Total	14,6	4,4	15,3	21,1

Tabla 1. Número medio de unidades por centros en varios niveles de enseñanza según el tamaño de los municipios. Elaboración propia a partir de Ministerio de Educación y Formación (2019)

Tamaño del municipio	% Población
Hasta 1.000	3,19
1.001 a 2.000	----
2.001 a 5.000	9,54
5.001 a 10.000	8,38
10.000 a 20.000	10,68
20.000 a 100.000	28,39
100.000 a 500.000	23,56
Más de 500.000	16,27
Total	100

Tabla 2. Número de habitantes y % de población representada según el tamaño de los municipios.
Elaboración propia a partir de los datos del Instituto Nacional de Estadística (2019)

Podemos estimar con estos cálculos que del orden del 40% de la población estará escolarizada en centros de una línea o dos líneas. Esto supone al respecto del posible diseño de experimentos en educación que los centros de línea 1 tendrían que ser desechados. Pasando a los de línea dos, es imposible elegir aleatoriamente los grupos de control ni los grupos de intervención para pruebas ligadas a un nivel curricular determinado, luego tendríamos una limitación global de difícil solución. La única solución sería formar grupos ficticios por extracción al azar de diferentes centros escolares. Pero estos centros deberían tener garantías de homogeneidad en muchas de las variables que se necesita tener controladas. Por otra parte, la variabilidad de los grupos así formados debería de ser la misma. Sin embargo, la pluralidad de centros implicaría pluralidad de intervenciones, e introduciríamos otra variable de difícil control. Minimizamos esas otras variables externas trabajando con los centros de línea 1 y línea 2, aunque existen las limitaciones anteriormente mencionadas en este párrafo para los diseños experimentales en estos centros. Además, incluso en ellos se manifiesta la dificultad de la elección al azar de un grupo de intervención y un grupo de control, existiendo también las variables ocultas o no controladas que implican una contingencia a cuantificar, así como la variabilidad de las muestras. Por lo tanto, supongamos que elegimos al azar varios centros de línea 2. Entonces nos centramos en un curso concreto: 5º de EP. Para este curso pensamos que podemos elegir al azar uno de los dos grupos de como grupo de control y el otro queda fijado como grupo de intervención, y creemos que se cumple una cierta aleatoriedad en la asignación del alumnado a los grupos. Al mantenernos en el mismo centro, tenemos la posibilidad de controlar los procesos de sesgo por la historia entre O1 y O2, o de maduración, y hasta podemos controlar el proceso de RTM realizando la misma prueba en varios centros con diferentes niveles. Parecería entonces que habríamos soslayado los problemas estructurales para construir un diseño experimental al que aplicar el análisis o contraste de hipótesis. Sin embargo, no es así. La razón estriba en la manera como se forman los grupos en cada centro. Este es uno de los problemas más graves que hay en los centros de línea 2 y línea 3. La repetición se puede dar en cualquier curso EP, dado que con la ley actual han desaparecido los ciclos, y en teoría existen herramientas administrativas que pueden posibilitar una cierta homogeneidad entre los grupos. Pero la realidad es que en los centros de línea dos, el alumnado repetidor, con necesidades especiales, con especial dificultad, conflictividad o carencias, se distribuye de manera notoriamente asimétrica entre unos grupos. Esto convertiría en casi imposible que un grupo cumpliera las condiciones estadísticas precisas para ser un buen grupo de control, y dificulta grandemente alcanzar un diseño experimental real.

Por otro lado, al respecto del control de las variables externas a la intervención, se identificarían los diferentes estilos didácticos, las diferencias en la evaluación, entre otros aspectos, que podrían influir notoriamente en los parámetros de variabilidad de los cursos. Los maestros de los distintos grupos y/o centros, conscientemente o no, podrían inducir unas diferencias en las medias, medianas, o desviaciones estándar de los grupos que hagan muy difícil conseguir las necesarias condiciones de homogeneidad, homocedasticidad y normalidad.

Por todo lo dicho anteriormente, este trabajo se llevó a cabo con un grupo de 5º curso de EP de un centro ordinario, donde se da la aleatoriedad en la asignación del alumnado a los grupos que se asume en un centro de sus características. La fiabilidad interna de las herramientas test (un pretest, un postest y una prueba externa elaborada por la tutora del curso), así como la normalidad de las variables resultantes de las pruebas, se basó en la teoría de los ítems y elementos tradicionales, calculándose los parámetros alfa de Cronbach, y sus aplicaciones subsidiarias a la valoración de ítems.

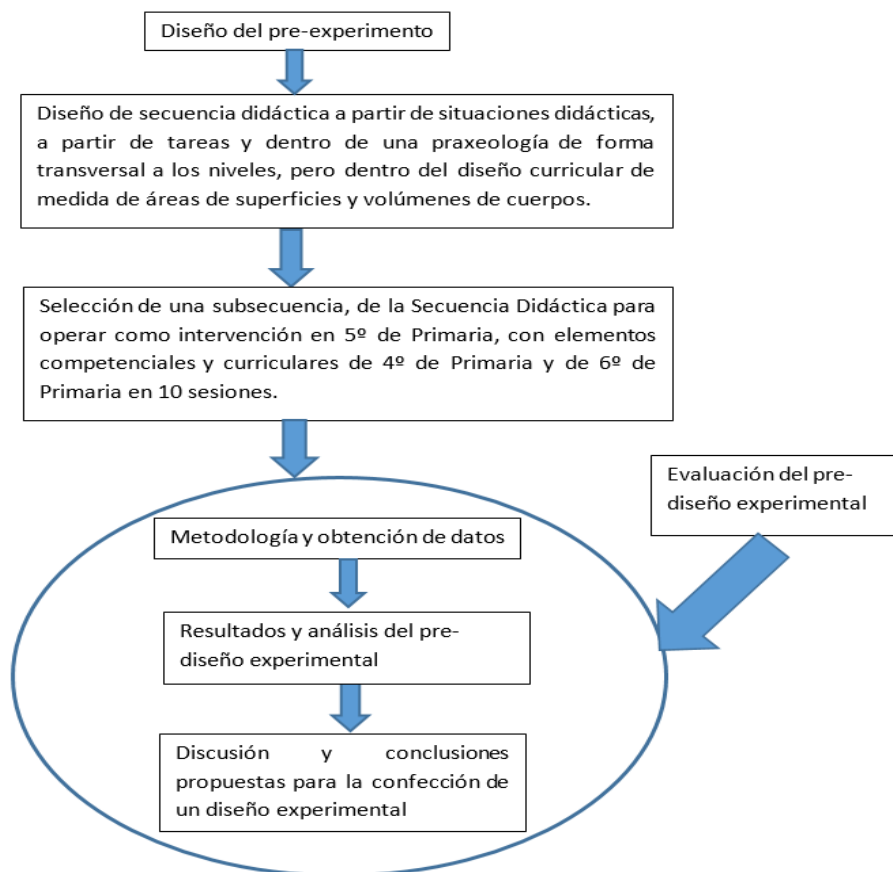


Figura 1. Esquema del diseño del pre-experimento y de los objetivos desarrollados en el presente trabajo

2. Metodología para la validación de la propuesta

Se validó de forma empírica el posible efecto de mejora en la comprensión de los conceptos de área y volumen mediante la intervención didáctica de una propuesta educativa que se encuentra en (Rieiro-Marín, Ocaña, García-Moya, y Fernández-César, 2019). Para ello, se siguieron los trabajos de García (2012) y Gullickson (2000).

Dado el funcionamiento de los centros educativos y las condiciones específicas del colegio en el que se desarrolló la investigación, se contemplaron las orientaciones metodológicas propuestas por Barlow y Hersen (1988) y Salas (2013).

Para dicha validación se realizó un diseño pre-experimental de estudio de caso único con un pretest (PRT-6, compuesto con 6 ítems) y un postest (PST-9, formado por 9 ítems, de los que 6 están en el PST-6 e implican actividades similares a las del PRT-6, a los que se añaden tres tareas de mayor dificultad), que se reforzó con una prueba realizada por la tutora para el estudio de consistencia externa, y se llamó último examen. La prueba externa fue diseñada por la tutora, y contenían tanto elementos del desarrollo curricular relacionados con la experiencia didáctica como con otros ajenos a la misma (véase la figura 2).

El grupo de intervención se constituyó por 24 estudiantes que cursaban 5º curso de Educación Primaria en el Colegio Nuestra Señora del Consuelo de Yuncos (Toledo).

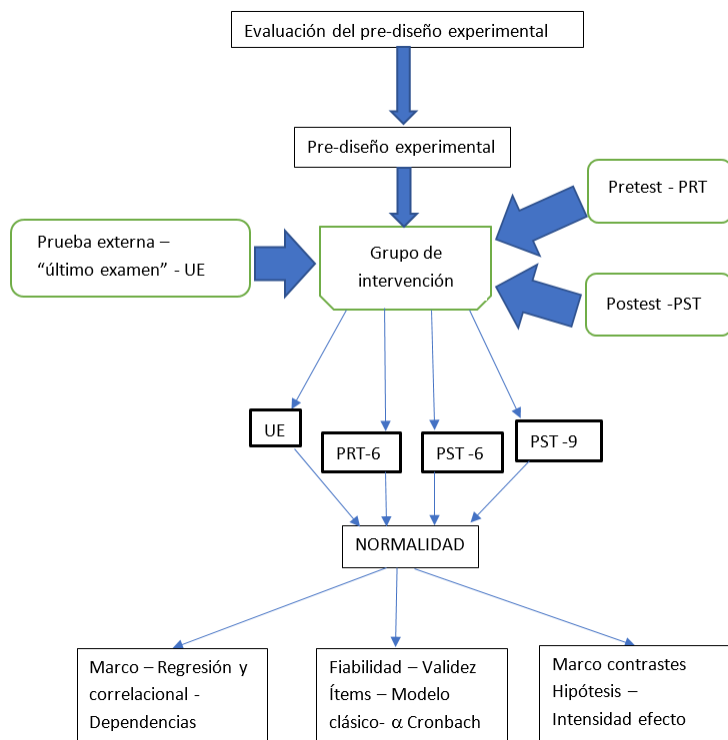


Figura 2. Esquema de la metodología empleada para la evaluación del diseño pre-experimental en el marco cuantitativo

2.1. Instrumento

El pretest y el posttest se elaboraron *ad hoc*, con ítems de respuesta múltiple, y cuatro opciones de respuesta entre las cuales solo una opción era correcta. De este hecho se informó a los alumnos.

El pretest se diseñó tanto para detectar los posibles errores conceptuales como los conocimientos iniciales de los escolares antes de la intervención didáctica, y estuvo compuesto de seis ítems. El posttest se diseñó para medir la posible mejora en dichos conocimientos, la persistencia o evolución de los errores tras la intervención, así como para abordar las posibles dificultades surgidas con los ítems propuestos. Este último se compuso por nueve ítems (véase la tabla 3).

	Pretest	Postest
Ítem 1	Conservación del área de una figura (a)	Conservación del área de una figura (a)
Ítem 2	Confusión perímetro-área (a)	Confusión perímetro-área (a)
Ítem 3	Error de localización de la altura de una figura plan (a)	Error de localización de la altura de una figura plan (a)
Ítem 4	Error de figura no pavimentada (a)	Error de figura no pavimentada (a)
Ítem 5	Conservación del volumen de una figura (v)	Conservación del volumen de una figura (v)
Ítem 6	Dominio espacial (v)	Dominio espacial (v)
Ítem 7		Conservación del área de figuras (a)
Ítem 8		Conservación perímetro-área (a)
Ítem 9		Dominio espacial (v)

Tabla 3. Tipos de errores que analizan los diferentes ítems sobre áreas (a) y volúmenes (v) en las pruebas diseñadas

La proporción de los ítems de áreas y volúmenes sobre el total se adecuó a la carga curricular de medida de superficies de 5º curso de Educación Primaria. Además, para poder hacer un análisis correlacional de los instrumentos, se relacionaron los seis primeros ítems del pretest y del postest, donde se variaron las figuras, pero se mantuvieron los errores que cada ítem analizaba, así como su marco competencial. Cabe destacar que las tres últimas preguntas del postest se incluyeron para evitar que las respuestas azarosas enturbiaran el resultado, y estas presentaron un más alto nivel de dificultad que las seis primeras.

Los ítems 1, 4 y 5 del pretest formaron parte de investigaciones lideradas por autores clásicos como Piaget, y se referenciaron en el trabajo de Del Olmo, Moreno y Gil (1989). El ítem número 6 se encuentra recogido en García Pérez (2007) y el resto fueron de creación propia (véase la figura 3). En la elaboración del postest se tomó el ítem número 4 del libro de Del Olmo, Moreno y Gil (1989), el número 6 de García Pérez (2007) y el número 8 del trabajo de Godino, Batanero y Roa (2002). El resto de los ítems fueron de creación propia, siguiendo estándares de la propuesta (véase la figura 4 donde aparecen ejemplos del pretest y postest y el [enlace](#) donde se muestran las pruebas completas).

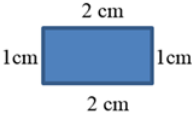
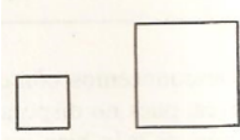
<p>1. Observa este rectángulo. ¿Cuál es su área?</p> <p>a) 6 cm^2 b) 2 cm^2 c) 3 cm^2 d) Ninguna de las anteriores</p>	<p>2. ¿Cuántos cuadrados pequeños hay en el grande? Señala con una cruz: 1 2 3 4 5</p>
	

Figura 3. Ítems del pretest para ejemplificación sobre confusión de perímetro y área (1) y errores de figuras no pavimentadas (2)


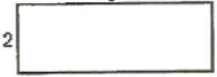

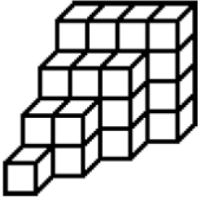
<p>1. ¿Cuál es el área de los siguientes rectángulos?</p> <p>A.  B.  C. </p> <p>A. _____ B. _____ C. _____</p>	<p>2. ¿Cuántos cubos hay en la figura en total? ¿y cuántos cubos no se ven?</p> <p></p> <p>a) Hay 25 cubos en total y no se ven 14 b) Hay 30 cubos en total y no se ven 14 c) Hay 31 cubos en total y no se ven 12 d) Hay 30 cubos en total y no se ven 12</p>
---	---

Figura 4. Ítems del post-test para ejemplificación sobre confusión de perímetro y área (1) y dominio espacial(2)

2.2. Análisis estadístico

El soporte estadístico de la metodología aplicada en el diseño del pre-experimento y el uso de paquetes estadísticos, se pueden encontrar en Vicente et al. (2005). Para el tratamiento de contrastes se siguió a Martín (2001), y para la adecuación de la estadística a la educación a Morales (2008), y Pérez et al. (2009).

Se empleó un procedimiento estadístico descriptivo, calculando los porcentajes de respuestas correctas por ítem, y las puntuaciones totales, tanto en el pretest como en el postest. Para la valoración de la intervención, se usó la diferencia entre el pretest y el postest.

Al ser un diseño pre-experimental se completó con una variable externa, la cual fue realizada por la tutora de los participantes, y se estudió la asociación entre el pretest, postest y el último examen. Para

ello, se empleó una escala de 10 puntos y las asociaciones entre las puntuaciones se realizaron mediante los respectivos coeficientes de correlación.

El procedimiento estadístico inferencial sobre las comparaciones de medias de las tres pruebas se realizó con contrastes paramétricos, previa comprobación de normalidad mediante tres tests: Kolmogorov-Smirnov con correlación de Lilliefors (KSL), sin corrección de Lilliefors (KS), y Shapiro-Wilk (SW). Dado el tamaño de la muestra ($N=24$), se consideró concluyente aquel resultado aportado por al menos dos de las tres pruebas de normalidad. Se emplearon los contrastes para muestras independientes o relacionadas valorando los coeficientes de correlación obtenidos para las puntuaciones totales de pretest, postest y el último examen.

Para afrontar el análisis de los resultados de las pruebas con herramientas de estadística inferencial, y realizar los contrastes de hipótesis entre las puntuaciones totales medias entre el pretest, postest y último examen, se estudió previamente la normalidad o no de dichas puntuaciones. Para ello se aplicó el criterio indicado, las puntuaciones totales para el pretest ($Z_{K-S} = .759$; $p = .612$; $Z_{K-S-L} = .155$; $p = .141$; $Z_{S-W} = .963$; $p = .511$), el postest con 9 ítems ($Z_{K-S} = .696$; $p = .717$; $Z_{K-S-L} = .142$; $p = .200$; $Z_{S-W} = .929$; $p = .091$), y el último examen ($Z_{K-S} = .673$; $p = .756$; $Z_{K-S-L} = .137$; $p = .200$; $Z_{S-W} = .908$; $p = .032$), las cuales fueron consideradas normales, por lo que se trabajó con ellas en el marco de la estadística paramétrica.

3. Resultados y discusión

El tratamiento estadístico de los datos se realizó con el programa SPSS, versión 15.0, así como con el programa Excel, y para la gestión general del procesamiento se siguieron las orientaciones de análisis en educación y área afines de Bisquerra (1987 y 1989).

En la figura 5 se muestran los porcentajes de respuestas correctas para los seis ítems que forman el pretest y postest (PST-6), y la diferencia con la que se pretende describir el posible efecto de la intervención. Los ítems 1, 2 y 5 presentan incrementos superiores al 40%, el 3 y el 6 muestran incrementos entre el 20% y el 30%, pero el ítem 4 sufre un decremento de aproximadamente el 15%, lo que indica una inadecuación entre el enunciado del pretest y el del postest. La mejora media del alumnado es próxima al 40%.

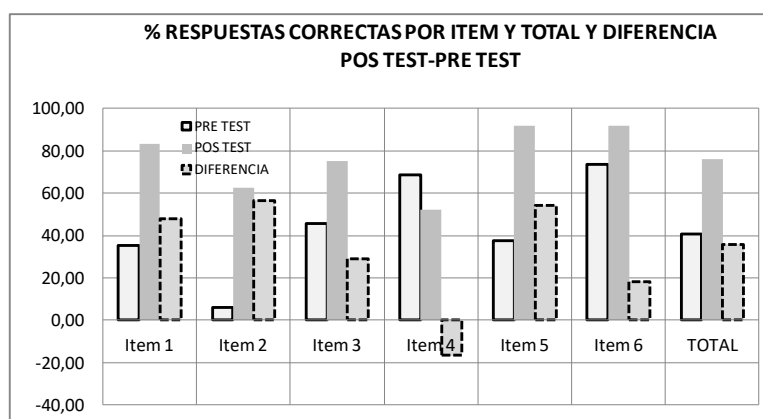


Figura 5. Porcentaje de respuestas correctas en el pretest y postest y diferencia entre ambos

El promedio de respuestas correctas en el pretest fue del 40%, mientras que en el postest con seis ítems estuvo próximo al 80%. Esto significa una mejora entorno al 40%. Con el postest de nueve ítems, el completo, el resultado fue próximo al 80%, pese a la mayor dificultad de los tres últimos ítems. La diferencia es, por tanto, muy similar, independientemente del postest considerado (véase la figura 6).

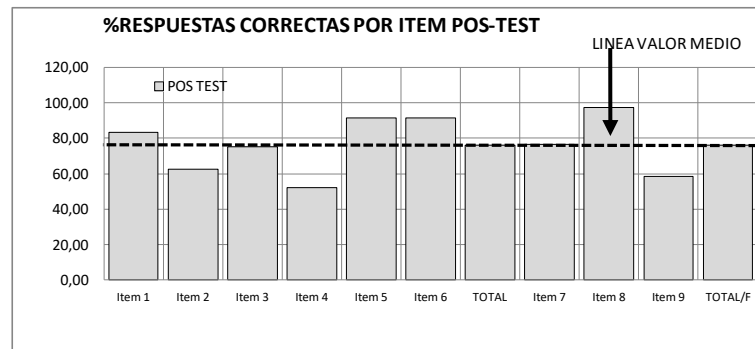


Figura 6. Porcentaje de respuestas correctas por ítem del postest

Si los resultados se muestran en un diagrama de cajas y bigotes (véase la figura 7), lo que posibilita la comparación de la puntuación total en sendos tests, se percibe la falta de intersección entre los recorridos de las dos cajas, lo que evidencia la desigualdad entre las medianas y las medias obtenidas por los alumnos en sendas pruebas.

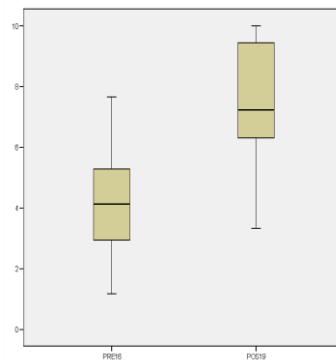


Figura 7. Diagrama de cajas y bigotes de las respuestas correctas totales en el pretest y el postest

Si se comparan las puntuaciones medias en el postest con las obtenidas en la prueba externa, se ve que son 7.59 (DT=1.91) y 6.25 (DT=2.14), respectivamente. Entre una nota y otra, existe una diferencia de 1.34 puntos. Además, el postest es suspendido por 1 alumno mientras que el último examen es suspendido por 4. Por otro lado, son similares las dispersiones de las puntuaciones obtenidas por los alumnos en ambas pruebas.

Los estadísticos básicos en las notas están en escalas de 10. Sin embargo, la escala para cada ítem no es 10, sino la que se puede ver en la tabla 4. En la misma se detalla la puntuación en el examen externo (último examen) en escala de 10. Además, los valores de los ítems se muestran en escala absoluta sobre los totales posibles definidos en la línea PUN.ITEM. Se observan valores medios por ítem en escala absoluta en escala de 10, y en porcentaje de respuestas correctas e incorrectas. Los totales por sujetos se han codificado en escala absoluta y en escala de 10.

Postes para determinar errores y superar dificultades

PUNT.ITEM	4	4	2	2	3	17	408	3	3	4	27	648		
SUJETOS	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Total	total nota	Ítem 7	Ítem 8	Ítem 9	Escala -Total	Total nota	Último exa
Alumno 1	4	4	2	2	2	3	17	10,00	3	3	2	25	9,26	5,5
Alumno 2	4	4	2	1	2	3	16	9,41	3	3	4	26	9,63	7,75
Alumno 3	4	4	2	2	2	3	17	10,00	3	3	4	27	10,00	8,5
Alumno 4	4	4	2	2	2	3	17	10,00	3	3	4	27	10,00	7,5
Alumno 5	0	4	2	2	2	3	13	7,65	1	2	2	18	6,67	5,75
Alumno 6	4	0	2	1	2	3	12	7,06	1	3	0	16	5,93	7,5
Alumno 7	4	0	0	1	2	3	10	5,88	3	2	0	15	5,56	1,5
Alumno 8	4	4	2	2	2	3	17	10,00	3	3	4	27	10,00	8,5
Alumno 9	4	0	0	0	2	3	9	5,29	3	3	4	19	7,04	6
Alumno 10	4	0	0	0	2	3	9	5,29	1	3	1	14	5,19	2,25
Alumno 11	4	4	2	0	2	3	15	8,82	3	3	0	21	7,78	8,25
Alumno 12	4	0	0	1	0	3	8	4,71	0	3	4	15	5,56	6
Alumno 13	4	0	2	0	2	3	11	6,47	3	3	1	18	6,67	2,25
Alumno 14	0	4	2	2	2	3	13	7,65	3	3	0	19	7,04	8,25
Alumno 15	4	0	0	0	2	3	9	5,29	3	3	4	19	7,04	4,25
Alumno 16	4	0	2	0	2	1	9	5,29	1	3	1	14	5,19	5
Alumno 17	4	4	2	1	2	3	16	9,41	3	3	0	22	8,15	8,25
Alumno 18	4	4	2	2	2	3	17	10,00	3	3	4	27	10,00	7
Alumno 19	0	0	2	0	2	1	5	2,94	0	3	1	9	3,33	5,5
Alumno 20	4	4	2	2	2	3	17	10,00	3	3	4	27	10,00	7
Alumno 21	4	4	0	2	2	3	15	8,82	3	3	4	25	9,26	9
Alumno 22	4	4	2	0	0	3	13	7,65	3	3	4	23	8,52	4,5
Alumno 23	0	4	2	2	2	3	13	7,65	0	3	4	20	7,41	8,25
Alumno 24	4	4	2	0	2	1	13	7,65	3	3	0	19	7,04	5,75
Alumno 25	80	60	36	25	44	66	311	7,62	55	70	56	492	7,59	6,25
Alumno 26	96	96	48	48	48	72	408	7,62	72	72	96	648	7,59	6,25
Alumno 27	0,83	0,63	0,75	0,52	0,92	0,92	0,762	MEDIAS	0,76	0,97	0,58	0,759	MEDIAS	MEDIAS
Alumno 28	83,33	62,50	75,00	52,08	91,67	91,67	76,23		76,39	97,22	58,33	75,93		
Alumno 29	83,33	62,50	75,00	52,08	91,67	91,67	76,23		76,39	97,22	58,33	75,93		
Alumno 30	16,67	37,50	25,00	47,92	8,33	8,33	23,77		23,61	2,78	41,67	24,07		
Alumno 31	3,33	2,50	1,50	1,04	1,83	2,75	12,96	7,62	2,29	2,92	2,33	20,50	7,59	6,25
Alumno 32	8,33	6,25	7,50	5,21	9,17	9,17	7,62		7,64	9,72	5,83	7,59		
Alumno 33	1,52	1,98	0,88	0,91	0,56	0,68	3,54	2,08	1,16	0,28	1,79	5,15	1,91	2,14

Tabla 4. Datos del postest por ítem y sujeto, con resultados descriptivos

En la tabla 5 se muestran escalas globales de puntuaciones en el pretest y postest sobre escala de 10. Se tabulan las medias, las desviaciones típicas y el coeficiente de variación de Pearson (CVP) en %. En la misma tabla, ITEM3POS designa al tercer ítem del postest, y de igual manera, ITEM5PRE designa al quinto ítem del pretest.

	PRE16	POS16	POS19	ITEM1POS	ITEM2POS	ITEM3POS	ITEM4POS	ITEM5POS	ITEM6POS	ITEM7POS	ITEM8POS	ITEM9POS	ITEM1PRE	ITEM2PRE	ITEM3PRE	ITEM4PRE	ITEM5PRE	ITEM6PRE	N válido (según lista)
N	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00	24,00
Media	4,07	7,62	7,59	8,33	6,25	7,50	5,21	9,17	9,17	7,64	9,72	5,83	3,54	0,63	4,58	6,88	3,75	7,36	
Desv. típ.	1,77	2,09	1,91	3,81	4,95	4,42	4,54	2,82	2,25	3,87	0,94	4,46	4,77	2,12	5,09	4,12	4,95	4,28	
CVP %	43,4884	27,4272	25,1027	45,6834	79,125	58,976	87,1587	30,7992	24,5824	50,6381	9,6698	76,5265	134,763	338,795	111,050	59,9439	131,87	58,1791	
	9	15	29	27	6	8	38	44	2	9	6	9	32	2	35	1	6	6	

Tabla 5. PRE16-Escala de pretest ítems 1 al 6; POS1-6-Escala de postest ítems 1 al 9, valores medios, desviaciones típicas y coeficientes de variación

El estudio de fiabilidad interna no se realiza para el pretest, dado los valores tan altos del coeficiente de variación de Pearson (CVP) que se obtienen (Tabla 5). Este supera el 300% en algún ítem, y en otros llega hasta el 100%, indicando que la prueba del pretest adolece de muy alta dispersión. Esto puede interpretarse como la consecuencia de una cumplimentación de este azarosa por parte de los alumnos, relacionada básicamente con la ausencia de conocimientos adquiridos. Este resultado es esperable, pues la prueba mide unos niveles de conocimientos de partida en el marco del experimento que no tenían por qué ser conocidos, o al menos no significativamente. No es el caso de los valores del CVP para el postest, que son del orden del 50%. Esto hace factible el estudio de la fiabilidad mediante el coeficiente alfa de Cronbach.

El valor del alfa de Cronbach, para el postest con 9 ítems es .564, próximo a .6. Hay razones que avalan la suficiencia de este valor en esta investigación. Autores como Nunnally (1967, p.226), sostienen que en marcos como el presente, pueden ser adecuados valores entre .5 y .6 del alfa de Cronbach (en otros marcos serían necesarios valores $\geq .7$). Sucede que el análisis de fiabilidad interna con este coeficiente está pensado para test con más ítems. En nuestro caso el número de ítems del test no está determinado por las necesidades de consistencia interna con este coeficiente, sino que está ajustado por las condiciones reales de una prueba para alumnos de 5º de EP. Al evaluar el alfa de Cronbach del postest con 6 ítems y con nueve, éste pasa de cerca de .5 a casi .6, y esto confirma que en nuestro caso los valores de este coeficiente están influidos por la cantidad de ítems de los que se compone el test. En este formato, un incremento de ítems notorio produciría valores más altos. Sin embargo, para nuestro objetivo didáctico es más adecuado mantener el tamaño de la prueba en su forma actual.

Los estadísticos para la fiabilidad de los ítems se pueden ver en la tabla 6. Ahí se constata que los ítems 1 y 5 deben ser modificados, pues el valor de correlación ítem-total es muy bajo para ambos, y el alfa de Cronbach mejora si se eliminan. Por ello de la información anterior se pueden esperar mejoras de la consistencia interna del test reformulando los ítems 1 y 5.

Media	Varianza	Desviación típica	Nº de elementos		
20,500	26,522	5,1499	9		
a)					
Ítem	Media de la escala si se elimina el elemento	Varianza de la escala si se elimina el elemento	Correlación elemento-total corregida	Correlación múltiple al cuadrado	Cronbach si se elimina el elemento
1	17,1667	23,275	,063	,517	,607
2	18,0000	14,609	,529	,723	,414
3	19,0000	24,522	,139	,518	,563
4	19,4583	21,650	,479	,663	,488
5	18,6667	26,319	-,020	,351	,582
6	17,7500	23,587	,378	,442	,524
7	18,2083	19,824	,518	,640	,458
8	17,5833	25,906	,187	,290	,563
9	18,1667	18,188	,265	,434	,545
b)					

Tabla 6. Estadísticos de escala del postest en puntuación total (a) y por ítem (b)

Las preguntas de respuesta múltiple tenían 4 opciones. Incluir más opciones de respuesta también contribuiría a incrementar el valor del alfa de Cronbach, sensible al número de opciones. Pero no es lo usual para estos alumnos y ese tipo de pruebas de conocimientos.

3.1 Marco Correlacional

En el marco correlacional, los valores del coeficiente de correlación de Pearson son suficientemente significativos para las relaciones entre el ítem2 y el ítem3 (.547(**)), entre el ítem 4 (.617(**)) y el ítem7 (.426(*)), así como entre el ítem1 y el ítem7 (.509(**)), y el ítem4 e ítem6 (.443(*)).

Ampliando el marco correlacional a las relaciones de correlación y regresión, entre las diferentes variables pretest, postest y último examen, podemos obtener una información complementaria, con cierta significación (véase el siguiente [enlace](#)).

En el marco del análisis correlacional entre el pretest, postest, y el último examen (véase la tabla 7), encontramos que no existe correlación significativa entre las puntuaciones totales del pretest y del

postest ($\rho = .043$; $p = .844$). Lo mismo sucede con el pretest y el último examen ($\rho = .041$; $p = .849$). Por lo tanto, la puntuación total en el pretest respecto de las otras dos pruebas debe ser considerada independiente. Sin embargo, la correlación entre el postest con 9 ítems y el último examen es alta y significativa ($\rho = .598$; $p = .02$), por lo que las puntuaciones totales en estas pruebas deben considerarse relacionadas.

La independencia entre los resultados del pretest y el postest permite afirmar que el incremento medio observado en el grupo posteriormente a la intervención no es debido a los niveles competenciales iniciales del alumnado, sino, salvando variables ocultas no controladas, al efecto de la intervención con la propuesta didáctica diseñada. Enfatizamos que este hecho se manifiesta tanto en la relación de los seis ítems concordantes, como en la existente en los nueve ítems del postest. De igual manera, podemos comprobar que los resultados del pretest no se correlacionan con los resultados del último examen, y que, por el contrario, son significativas las correlaciones entre los resultados del postest con seis ítems, postest con nueve ítems, y el último examen. Los valores de la correlación con el control externo del último examen se encuentran próximos a .6, valor similar al del coeficiente alfa de Cronbach. Aunque este valor muestre una correlación débil, el hecho de ser entre nuestro postest y una prueba externa, nos aportan una cierta garantía de robustez de la prueba.

Coeficiente de correlación de Pearson	N	Correlación	Sig.
Par 1 TOT_NOTA_PRE16 y TOT_NOTA_POS16			
Par 2 TOT_NOTA_PRE16 y TOT_NOTA_POS19	24	0,237	0,265
Par 3 TOT_NOTA_PRE16 y TOT_NOTA_ULTIMO_EXAMEN	24	0,043	0,844
Par 4 TOT_NOTA_POS16 y TOT_NOTA_POS19	24	0,041	0,849
Par 5 TOT_NOTA_POS16 y TOT_NOTA_ULTIMO_EXAMEN	24	0,925	0
Par 6 TOT_NOTA_POS19 y TOT_NOTA_ULTIMO_EXAMEN	24	0,598	0,002
	24	0,553	0,005

Tabla 7. Coeficiente de correlación de Pearson entre resultados significativos del estudio

En la Figura 8 se observan las regresiones entre las diferentes pruebas. Así, en 8 (a) se percibe que no hay ninguna posibilidad de relación entre las medidas del pretest y la evaluación de la marcha ordinaria de los alumnos. Esto puede ser debido a dos causas: una, a que no existen apenas elementos de geometría del tema abordado en la intervención, y que el último examen recoge preguntas sobre otros aspectos tratados en el trimestre; y otra, que la herramienta del pretest y la del último examen, eran notoriamente diferentes en su confección. Posiblemente la no existencia de relación se deba a las dos razones. Esto, junto a lo anteriormente analizado, nos dice que el pretest solo recoge los conocimientos "sociales" de los alumnos sobre el tema de medida de áreas de superficies y volúmenes de cuerpos, y que los aprendizajes en el curso previo no eran significativos por la intensidad de los mismos, o por el carácter introductorio con el que se tratan dichos temas en ese curso en la mayor parte de los centros. La regresión entre el postest y el último examen, 8 (b), solo explica un 35,78 % de la varianza lo que no es nada significativo, pero si muestra una tendencia. En el caso del postest completo de 9 ítems, figura 8(c), solo explica el 30,56 % de la varianza, lo que sigue siendo un valor pequeño. No obstante, el último examen y el postest tienen una relación no nula, pues existen elementos de covariación conjunta, aunque dicha covariación no tiene por qué implicar relación de causa – efecto únicamente. Lo que es muy significativo es lo que se puede ver en las figuras 8 (d) y 8(e), que ratifica la nula influencia entre pretest y postest. Esta ausencia de relación entre pretest y postest, junto con el control "histórico", entre pretest y postest, nos permite valorar que los demás elementos de disminución de fiabilidad no están presentes, y la atribución causal al resultado del postest pueden ser atribuidos a la intervención X. Sin

embargo, una cuestión que no podemos resolver es la comparación con un grupo que haya abordado los mismos temas, elementos de dificultad y errores, con una metodología tradicional o clásica, de manera que pudiéramos atribuir comparativamente el efecto positivo y su intensidad al factor metodológico. Tampoco es susceptible esa afirmación, ni aunque fuese el grupo de comparación homogéneo, y fuese evidente la diferencia entre los incrementos de los resultados en el caso de la metodología clásica y la estructura del diseño de la secuencia didáctica, debido a que en los niveles de EP existen factores atribuibles al docente, como la empatía, el estímulo por la materia y otros aspectos emocionales, que podrían producir una influencia mayor que los aspectos técnicos, tecnológicos o metodológicos.

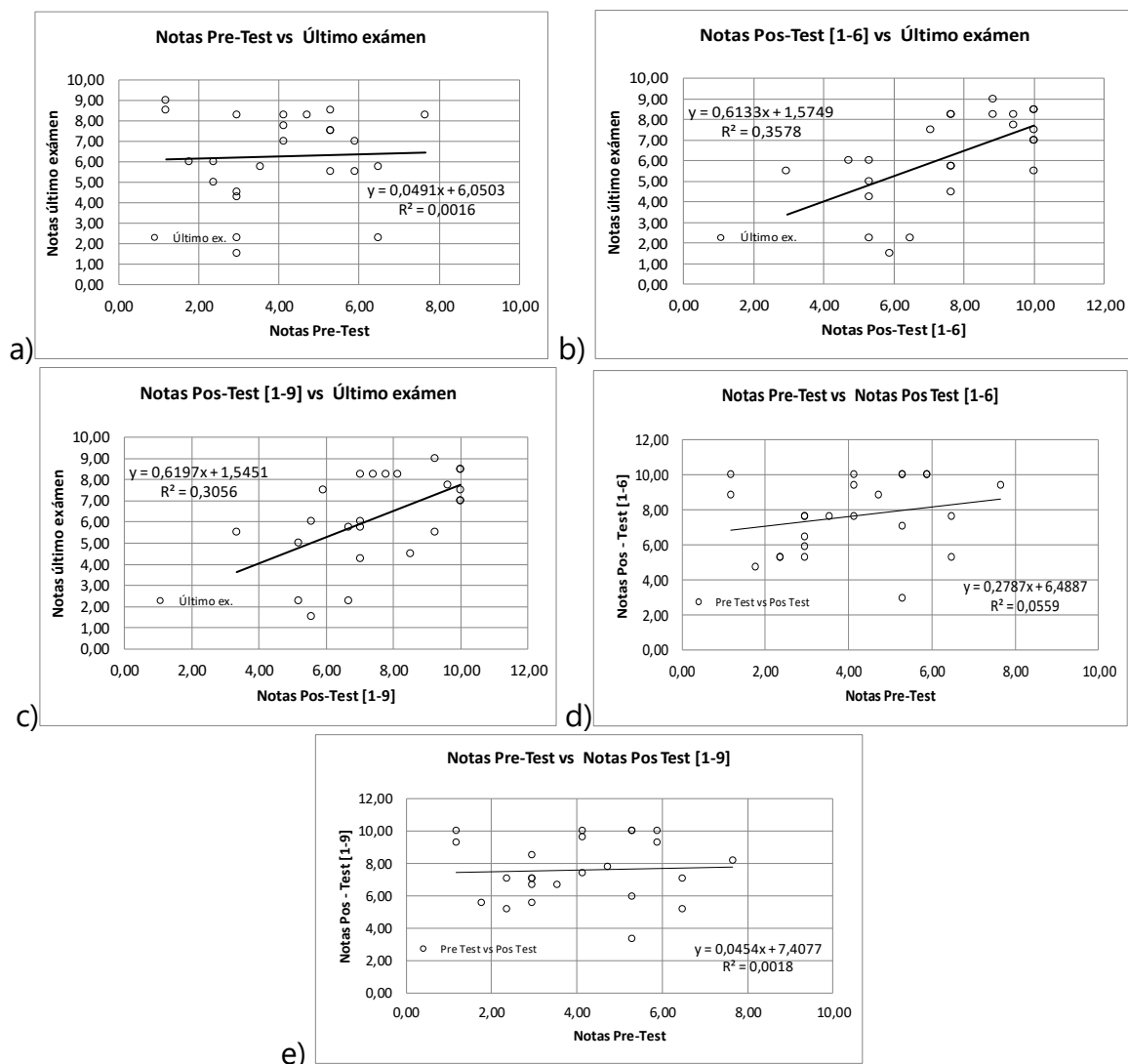


Figura 8. Diagramas de regresión lineal para diferentes variables; a) Regresión UE-vs-PR6 b) Regresión UE-vs-PS6 c) Regresión UE-vs-PS9 d) Regresión PS6 – vs- PR6 e) Regresión PS9-v-PR6

3.2. Contraste de medias

En el inicio de esta sección mostramos las diferencias entre las puntuaciones totales en el pretest y el postest. Aplicando ahora la inferencia estadística al contraste de las puntuaciones totales medias tanto con 6 ($t = -6.645$; $gl=45.741$; $p < .00$) como con 9 ítems en el postest ($t = -5.441$; $gl=44.95$; $p < .00$), confirmamos que estas son estadísticamente diferentes. Se pone de manifiesto una puntuación significativamente diferente y mayor en el postest, sin que pueda ser esto atribuido a las competencias

iniciales de los alumnos. Lo mismo se pone de manifiesto al comparar el pretest con el último examen ($t = -3.332$; $gl=45.98$; $p=.002$).

Aplicamos el contraste de medias a las puntuaciones en el posttest, tanto con 6 como con 9 ítems, frente a la prueba externa para ratificar la efectividad de la propuesta didáctica, y la robustez del método de medida de dicha efectividad. Para ello, comparamos las medias mediante un test t-student con muestras relacionadas, el cual revela que tanto el posttest con 6 ítems y el último examen ($t=3,550$; $gl=23$; $p=.002$) como el posttest con 9 ítems y esa misma prueba externa ($t=3,424$; $gl=23$; $p=.002$), aportan resultados estadísticamente diferentes. Por otro lado, como estas puntuaciones mostraban correlación media-alta y significativa, podemos atribuir la diferencia observada en las puntuaciones de los alumnos a la diferencia de competencias que demuestran tener en cada uno de los momentos en los que se administran las pruebas, correspondiendo el más alto al momento tras la intervención.

3.3. Valoración de la formulación de los ítems

Queremos valorar lo que aporta nuestra propuesta didáctica y evaluar si contribuye a eliminar errores de los alumnos sobre el área y el volumen, y superar las dificultades de aprendizaje que estos conceptos y su didáctica suponen. Para ello mostramos la tabla 8, en la que se ven las dificultades de aprendizaje que cada uno de los ítems del pre y el posttest pretende paliar. Se informa de si los porcentajes de respuesta correcta en valor absoluto en el pretest y en el posttest, son bajos, muy bajos, altos, muy altos, normales o anómalos, para cada ítem. Se emplea el mismo código para el incremento, y se añade en la última columna de la derecha la valoración estadística sobre el comportamiento didáctico medido por las diferencias entre los valores absolutos mencionados, en cuanto que sean ítems con comportamiento anómalo, significativo normal o altamente significativo.

Ítem	Pretest	Posttest	%Res.Correctas. Pre	%Res.Correctas. Pos	Incre:Pos-Pre	Valoración
1	Conservación del área de una figura (a)	Conservación del área de una figura (a)	Valor>30%(n)	Valor>80%(m.a.)	Incre>40%(a)	Altamente +
2	Confusión perímetro-área (a)	Confusión perímetro-área(a)	Valor>5%(b)	Valor>60%(a)	Incre>50%	Altamente +
3	Error de localización de la altura de una figura plan(a)	Error de localización de la altura de una figura plan(a)	Valor>40%(n)	Valor>70%(a)	Incre>20%(n)	Normal
4	Error de figura no pavimentada(a)	Error de figura no pavimentada(a)	Valor>70%	Valor>50%(b)	Incre>10%(n)	Anómalo
5	Conservación del volumen de una figura(v)	Conservación del volumen de una figura(v)	Valor>30%(n)	Valor>90%(m.a.)	Incre>50%(a)	Altamente +
6	Dominio espacial (v)	Dominio espacial (v)	Valor>70%	Valor>90%(m.a.)	Incre>15%(n)	Altamente +
7		Conservación del área de figuras (a)		Valor>70%(a)		Normal +
8		Conservación perímetro-área(a)		Valor>90%(m.a.)		Altamente +
9		Dominio espacial (v)		Valor>50%(b)		Normal +

Legenda = (a) =área / (v)=volumen / (n)=normal / (a)=alto/ (m.a.)=muy alto / (m.b.)=muy bajo

Tabla 8. Valoración de la evolución del porcentaje de respuestas correctas por ítem entre el pretest y el posttest

Los alumnos parten de unos errores y dificultades relacionadas con la conservación de áreas de figuras planas (ítem 1) y con la confusión del perímetro con el área (ítem 2), así como con la localización de la altura de una figura plana (ítem 3); también muestran errores en la conservación del volumen de un cuerpo (ítem 4). En el otro extremo encontramos el dominio espacial en las relaciones de volúmenes con cuerpos (ítem 6, 9) que parece estar altamente dominado por los alumnos de la muestra.

4. Conclusiones, limitaciones y prospectiva

En primer lugar, se concluye la eficacia real de esta intervención didáctica con alumnos de una clase de 5º de primaria, usando pretest y postest y evaluación externa de la maestra regular para la asignatura de Matemáticas. Se determinan niveles de consistencia interna del postest aceptables, y se ha analizado la calidad de los ítems, cuestionando solamente uno de ellos. Se constata la evidencia estadísticamente significativa de la mejora producida por la intervención sin relación causal con los niveles competenciales previos de los alumnos, ratificado de forma correlacional con una variable externa.

Para este diseño no podemos valorar efectos de sesgo no históricos al no haber grupo de comparación. Se puede inferir, que los materiales didácticos manipulables y visuales utilizados en la propuesta han contribuido al éxito de los resultados de los alumnos. Se ha probado la eficacia indistintamente para el tratamiento de áreas de superficies planas y volúmenes de cuerpos. Esto es así por la valoración alcanzada en los seis primeros ítems del postest, relacionados con los de pretest, y cuyo incremento no es cuestionable, además de por el contraste con los tres ítems de mayor dificultad competencial el del postest.

En segundo lugar, se ha podido comprobar, tras la búsqueda bibliográfica, que la mayoría de los estudios de caso único y propuestas publicadas sobre este tema, no tienen un análisis estadístico de resultados como el que metodológicamente se ha aplicado en este trabajo. Por lo tanto, se puede decir que este trabajo aporta alguna completitud y novedad metodológica en un tema carencial como es el de los diseños pre-experimentales.

Si bien es cierto que esta investigación inicial es mejorable, puede resultar de ayuda o base para la práctica de los docentes de Educación Primaria a la hora de impartir los contenidos de áreas de superficies y volúmenes de cuerpos de una manera más significativa que sirviéndose de la metodología tradicional. La propuesta fundamental de mejora pasa por realizar un incremento en el número de ítems de las dos pruebas al menos a 15 ítems, y trabajar con dos grupos de intervención y dos grupos de control, con dos intervenciones seriadas en el tiempo y de diferente dificultad a lo largo de un curso, con tiempos de respuesta y olvido suficientes. En su defecto, podría orientarse la subsiguiente investigación en un proceso colaborativo de construcción de un meta-análisis de diseños pre-experimentales de estudio de caso único. Se encuentran meta-análisis de este tipo en los trabajos de Letón y Pedromingo (2001), Botella Ausina y Gambara (2012), Botella Ausina, y Sánchez Meca (2015).

Para posibilitar la reproducción y falsación del trabajo, y el uso de nuestros datos para realizar un meta-análisis a quienes lo deseen, se han subido a la red todos los datos los materiales, los análisis estadísticos, y los resultados en el siguiente [enlace](#). Los resultados estadísticos de los análisis realizados en SPSS se pueden consultar en el siguiente [enlace](#).

Agradecimientos

Agradecemos al Colegio Nuestra Señora del Consuelo de Yuncos (Toledo) y a Rosa María Redondo Magán, tutora de 5º curso de Educación Primaria, su apoyo y colaboración en la aplicación de la propuesta de secuencia didáctica. Agradecemos a la exalumna de Grado de la Facultad de Educación de Toledo, D^a Paloma Ocaña Aranda, su implicación en el desarrollo de un Trabajo Final de Grado en el ámbito de iniciación a la Investigación Educativa en Primaria, su dedicación y la superación de numerosas dificultades.

Referencias

- Barlow, D., H., y Hersen, M. (1988). *Diseños experimentales de caso único. Estrategias para el estudio del cambio conductual*. Barcelona, España: Ediciones Martínez Roca.
- Bisquerra, R. (1987). *Introducción a la Estadística Aplicada a la Investigación Educativa. Un enfoque informático con los paquetes BMDP y SPSSX*. Barcelona, España: Ediciones PPU. S.A.
- Bisquerra, R. (1989). *Introducción conceptual al Análisis Multivariable. Un enfoque informático con los paquetes, SPSS-X, BMDP, LISREL y SPAD*. Barcelona, España: Ediciones PPU. S.A.
- Botella Ausina, J., y Gambará, H. (2012). *Qué es el Meta-Análisis*. Madrid, España: Grupo Editorial Siglo XXI, S.A.
- Botella Ausina, J., y Sánchez Meca, J. (2015). *Meta-Análisis en Ciencias Sociales y de la Salud*. Madrid, España: Editorial Síntesis.
- Brousseau, G. (1986). Fondements et méthodes de la didactiques des mathématiques. *Recherches en Didactique des Mathématiques*, 7(2), 33-115.
- Campbell Donald T., y Stanley, J.C. (1963). *Experimental and Quasi-Experimental Design for Research*. Boston, Massachusetts: Houghton Mifflin Company.
- Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques*, 19(2), 221-266.
- Chevallard, Y., Bosch, M., y Kim, S. (Febrero de 2015). *What is a theory according to the anthropological theory of the didactic?* CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Charles University in Prague, Faculty of Education; ERME. Prague, Czech Republic.
- Del Olmo, M.A., Moreno, M.F., y Gil, F. (1993). *Superficie y Volumen ¿algo más que el trabajo con fórmulas?* Madrid, España: Síntesis Educación.
- Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological*
- García Pérez, L. (2007). *Desarrollo de habilidades espaciales a través del uso de materiales concretos en niños de sexto grado de educación primaria*. Tesis de Maestría, Universidad Pedagógica Nacional, México. Recuperado de <http://www.remeri.org.mx/portal/REMERI.jsp?id=oai:tesis.upn.mx:24003>
- García Ramos, J. M. (2012). *Fundamentos Pedagógicos de la Evaluación. Guía Práctica para Educadores*. Madrid, España: Editorial Síntesis, S.A.
- Godino, J.D., Batanero, C., y Roa, R. (2002). *Medida de magnitudes y su Didáctica para Maestros*. Proyecto Edumat-Maestros. Universidad de Granada.
- Gullickson, R. A. (2000). *The need for student evaluation standards*. The Joint Committee on Standards for Educational Evaluation. *Institute of Great Britain and Ireland*, 15, 246-263.
- Instituto Nacional de Estadística (2019). *España en cifras 2019*. Recuperado de https://www.ine.es/prodyser/espa_cifras/2019/50/
- Letón, E., y Pedromingo M, A. (2001). *Introducción al Análisis de Datos en Meta-Análisis*. Madrid, España: Ediciones Díaz de Santos, S.A.
- Marsden, E., y Torgerson Carole, J. (2012). Single group, pre – and post – test research designs: Some methodological concerns. *Oxford Review of Education*, 38(5), 583-616.
- Martín, Q. (2001). *Contrastes de Hipótesis*. Madrid, España: Editorial la Muralla S.A.
- Ministerio de Educación y Formación (2019). Datos y cifras de la Educación en España, curso 2019-2020. Recuperado de <http://www.educacionyfp.gob.es/servicios-al-ciudadano/estadisticas/indicadores-publicaciones-sintesis/datos-cifras.html>
- Morales Vallejo, P. (2012). *Estadística aplicada a las ciencias sociales*. Universidad Pontificia Comillas. Madrid.
- Ortiz Aguilar, W, Enrique Hevia, F.M., y García La Rosa, J.E. (2018). Pre-Experimento para constatar cambios al aplicar un resultado científico. Ejemplo práctico de una metodología para mejorar el desarrollo de habilidades del pensamiento geométrico espacial. *Opuntia Brava*, 10(13), 329-326.
- Pérez, R., García L, J.L., Gil Pascual, J.A., y Galán, A. (2009). *Estadística Aplicada a la Educación*. UNED: Pearson. Madrid, España: Prentice Hall, S. A.
- Rieiro-Marín, I., Ocaña Aranda, P. O., García-Moya, M., & Fernández-César, R. (2019). Didactic proposal to overcome the difficulties in the learning of Area and Volume in Spanish Primary Education students. *Journal of Research in Science, Mathematics and Technology Education*, 2(3), 151-178. <https://dx.doi.org/10.31756/jrsmt.232>
- Rojas Álvarez, C.J. (2013). La instrucción geométrica y la representación plana de módulos multicubos en un grupo de alumnos: un diseño preexperimental. *Revista del Instituto de Estudios en Educación Universidad del Norte*, 19, 56-62.
- Romero, I. (29 de septiembre de 2012). *Dos propuestas metodológicas para la investigación en el aula: investigación-acción e investigación de diseño*. Conferencia presentada en Ciclo de conferencias en Educación Matemática de Gemad. Bogotá.

- Salas, E. (2013). Diseños pre-experimentales en psicología y educación: una revisión conceptual. *LIBERABIT*, 19(1), 133-141.
- Vicente, M.L., Girón, P., Nieto, C., y Pérez, T. (2005). *Diseño de Experimentos. Soluciones con SAS y SPSS*. Madrid, España: Pearson Prentice Hall.

Ignacio Rieiro Marín, profesor Titular de Matemáticas, Área de Didáctica, en la Universidad de Castilla La Mancha, Facultad de Educación de Toledo. Actualmente es profesor Colaborador Honorífico. Ha investigado en los siguientes temas en el ámbito de la Didáctica de las Matemáticas y la educación: Diseño de Evaluación para la Evaluación Diagnóstica y Competencial en Matemáticas en E. Primaria, Diseño de propuestas didácticas para la superación de dificultades y errores tipificados en diferentes áreas de la Matemáticas en E. Primaria, Conocimiento y deficiencias de los futuros maestros de Primaria en Didáctica de las Matemáticas, Confección de patrones e índices de rendimientos educativos no universitarios en España desde 1939 hasta la actualidad, La influencia del análisis de grabaciones de video en la mejora de la docencia.

Email: Ignacio.Rieiro@uclm.es

Melody García-Moya, investigadora en formación dentro del Doctorado en investigación en Humanidades, Artes y Educación de la Universidad de Castilla La Mancha. Colaboradora en el Grupo de investigación de Enseñanza de las Matemáticas (EMAT). Sus líneas de investigación son la resolución de problemas en Educación Primaria, y la enseñanza-aprendizaje de las Matemáticas en alumnado con necesidades específicas de apoyo educativo.

Orcid: <https://orcid.org/0000-0002-9634-5147>

Email: melody.garcia@uclm.es

Paloma Ocaña Aranda, Graduada en Maestro de Educación Primaria en la Facultad de Educación de Toledo, Universidad de Castilla La Mancha.

Email: palomaocara92@hotmail.com

Raquel Fernández César, profesora en el Departamento de Matemáticas, Área de Didáctica de la Matemática, en la Universidad de Castilla La Mancha, Facultad de Educación de Toledo. Es miembro del grupo de investigación "Mirada Crítica". Su campo de investigación es el dominio afectivo en la enseñanza-aprendizaje de las materias STEM, así como la exploración de metodologías activas centradas en el estudiante.

Orcid: <https://orcid.org/0000-0002-9013-7734>

Email: Raquel.FCezar@uclm.es