

CALIDAD DIDÁCTICA EN MATEMÁTICAS: UNA PROPUESTA DE EVALUACIÓN

LEÓN, Jaime

MEDINA-GARRIDO, Elena

NÚÑEZ, Juan L.

Universidad de Las Palmas de Gran Canaria

Las Palmas de Gran Canaria (España)

Jaime.leon@ulpgc.es

Resumen

Desgraciadamente el rendimiento y la implicación en las Matemáticas disminuyen durante la etapa de secundaria, por lo que los investigadores se encuentran ante el reto de encontrar soluciones a este problema. Para ello, llevamos a cabo dos estudios. En el primero, desarrollamos un instrumento para evaluar la calidad didáctica, concretamente evaluamos nueve factores: promover la utilidad y el interés, identificar y comprender los sentimientos negativos de los alumnos, promover la participación en clase, usar un lenguaje no controlador, diseñar actividades con un nivel de dificultad adecuado, centrarse en el proceso, estructurar las clases, feedback y, por último, ser afectuoso y cuidar a los alumnos. En el segundo estudio, observamos, a nivel grupal, que la calidad didáctica predecía las notas de Matemáticas a través de la implicación, y a nivel individual, observamos que la implicación predecía las notas de Matemáticas. Con esta investigación, no solo aportamos un instrumento para evaluar la calidad didáctica de forma fiable y válida, sino que también es una herramienta que puede utilizarse para desarrollar intervenciones que pueden impulsar a los alumnos a implicarse más en clase y, con ello, mejorar su rendimiento académico.

Abstract

Math achievement and engagement declines in secondary education; therefore, educators are faced with the challenge of engaging students to avoid school failure. Within self-determination theory, we address the need to assess comprehensively student perceptions of teaching quality that predict engagement and achievement. In study one we developed a questionnaire to assess nine factors: teaching for relevance, acknowledge negative feelings, participation encouragement, controlling language, optimal challenge, focus on the process, class structure, positive feedback, and caring. In the second study, at the classroom level, teaching quality was a predictor of behavioral engagement. Higher grades were observed in classes where students, as a whole, displayed more behavioral engagement. At the within level, behavioral engagement was associated with achievement. We not only provide a reliable and valid method to assess teaching quality, but also interventions could be designed based on the scale items to encourage students to persist and display more engagement on school duties, which in turn bolsters student achievement.

Palabras clave

Calidad didáctica; Evaluación del profesor; Matemáticas; Rendimiento académico;

Keywords

Academic Achievement; Teacher Quality; Teacher Evaluation; Mathematics Achievement

INTRODUCCIÓN

El rendimiento y la implicación de los estudiantes disminuye durante la etapa de educación secundaria (Kiemer, Gröschner, Pehmer, & Seidel, 2015). Es por ello que, durante esta etapa, los profesores e investigadores se enfrentan al reto de optimizar el potencial de los alumnos para que logren sus objetivos académicos. ¿Cómo hacerlo? Guiándonos por la teoría de la autodeterminación (TAD; Deci & Ryan, 1985, 2000), un marco teórico que explica el comportamiento y aprendizaje humano. De forma resumida y si nos centramos en el entorno educativo, la TAD propone que en la medida en que un alumno se siente competente, autónomo y vinculado a su profesor, será capaz de optimizar su potencial académico (Liu, Wang, & Ryan, 2016). Dentro de la TAD investigadores han observado evidencias sobre el papel que juega la calidad didáctica del profesor en el rendimiento de los estudiantes (Núñez & León, 2015). Sin embargo, no se conocen cuáles son los comportamientos específicos y concretos que caracterizan una calidad didáctica adecuada, por eso, para diseñar intervenciones eficaces debe ser una prioridad identificar los mecanismos que pueden usar los docentes para incrementar el rendimiento e implicación de sus estudiantes (Stroet, Opdenakker, & Minnaert, 2013, 2015).

Objetivos, preguntas de investigación e hipótesis

Para alcanzar este objetivo se realizaron dos estudios. En el primero, se diseñó una escala para evaluar la calidad didáctica; y, en el segundo, se probó si la escala predecía el rendimiento en Matemáticas a través de la implicación.

En el primer estudio pretendemos responder a la pregunta: ¿Se puede diseñar una escala fiable y válida con la que los alumnos evalúen la calidad didáctica de su profesor?

Nuestra hipótesis general en este primer estudio es que sí, y lo podemos formular con más precisión:

- Hipótesis 1a: se puede elaborar una escala que muestre adecuadas propiedades psicométricas que evalúen la calidad didáctica del profesor.
- Hipótesis 1b: los factores mostrarán relaciones significativas entre ellos.

En el segundo estudio nos centramos en analizar si la calidad didáctica del profesor de Matemáticas, evaluada mediante la escala desarrollada en el estudio anterior, predice el rendimiento de los estudiantes a través de la implicación. En este caso, antes de entrar en la formulación de las hipótesis de este segundo estudio, planteamos dos preguntas de carácter general: ¿Puede la calidad didáctica predecir la implicación de los alumnos? ¿La implicación predice las calificaciones de los alumnos?

La hipótesis general, por tanto, en este segundo estudio establece la relación entre calidad y rendimiento en los siguientes términos:

- Hipótesis 2a: la calidad didáctica del profesor se relaciona positivamente con la implicación de los alumnos de los alumnos.
- Hipótesis 2b: la implicación de los alumnos se relaciona positivamente con las notas de Matemáticas.

ESTUDIO 1

MÉTODO

Participantes

En este estudio participaron 548 alumnos de educación secundaria (52% chicos) con una edad media de 14.247 años ($DT = 1.123$). Los alumnos pertenecían a 24 grupos, con una media de 22.37 alumnos por grupo (mínimo = 14; máximo = 30, $DT = 3.70$). Los alumnos estaban entre 2º y 4º de Educación Secundaria Obligatoria, (2º de ESO, $n = 262$, $M_{edad} = 13.48$; 3º de ESO, $n = 124$, $M_{edad} = 14.38$; 4º de ESO, $n = 157$, $M_{edad} = 15.43$). Los institutos que formaron parte del estudio son de naturaleza variada: zonas rurales, periféricas y urbanas y los estudiantes son, en su gran mayoría, de familias de clase media. Durante la etapa de educación secundaria los alumnos cuentan con cuatro horas semanales de Matemáticas. La evaluación comenzó seis meses

después de que comenzara el curso, por lo que los alumnos habían tenido tiempo suficiente para saber cómo era el comportamiento de su profesor en el aula.

Participantes

El primer paso consistió en contactar con los departamentos de Matemáticas de institutos y contarles el proyecto para ver si querían participar. Finalmente, cuatro institutos formaron parte de la investigación. A la hora de realizar la evaluación, los alumnos dieron su consentimiento para formar parte del estudio y su participación fue estrictamente voluntaria y confidencial. Menos del 1% rechazó formar parte del proceso de evaluación. En la escala se explicaba a los alumnos que debían responder a las preguntas pensando exclusivamente en su profesor de Matemáticas. Los investigadores entregaron el cuestionario en marzo de 2015, y durante la recolección de datos entregaron a los alumnos las instrucciones, dieron las aclaraciones necesarias para completar la escala y se pidió al profesor que saliera del aula.

Instrumentos de medida

La escala fue diseñada por un grupo de expertos en la teoría TAD y se contó con asesoramiento por parte de profesores de Matemáticas. Se diseñó un conjunto de 83 ítems que abarcaba los distintos comportamientos de los profesores agrupados en factores. Concretamente, todos los ítems estaban clasificados dentro los siguientes factores: 1) Explicar la Utilidad de los Contenidos, 2) Cultivar los Intereses de los Alumnos, 3) Ofrecer Múltiples Opciones, 4) Identificar y Comprender los Sentimientos Negativos de los Alumnos, 5) Promover la Participación en Clase, 6) Usar un Lenguaje No Controlador, 7) Diseñar Actividades con un Nivel de Dificultad Adecuado, 8) Centrarse en el Proceso, 9) Dar Instrucciones Paso a Paso, 10) Preparar las Clases, 11) Feedback Rápido, 12) Feedback Específico, 13) Feedback sobre la Autorregulación, y 14) Ser Afectuoso y Cuidar a los Alumnos. Las respuestas se evaluaron de acuerdo a una escala Likert de 1 (*muy en desacuerdo*) a 7 (*muy de acuerdo*) puntos.

Análisis de datos

Para depurar la escala, se recurrió a diferentes técnicas estadísticas para determinar la estructura factorial óptima: análisis paralelo (Horn, 1965), modelo de ecuaciones estructurales exploratorio (ESEM, por sus siglas en inglés; Asparouhov & Muthén, 2009), modelo de ecuaciones estructurales bayesianas (BSEM, por sus siglas en inglés; Muthén & Asparouhov, 2012). En primer lugar, como recomiendan Asparouhov, Muthén, y Morin (2015), realizamos un BSEM, en el que los pesos factoriales en los factores no correspondientes (*crossloadings*) no tenía un valor de 0, sino que seguían una distribución de media 0 y varianza .01. A continuación, realizamos un ESEM con todos los ítems; sin embargo, los datos que obtuvimos no fueron esclarecedores porque no podíamos deducir el número de factores, ni podíamos saber a qué factor pertenecían muchos de los ítems. Esto nos llevó a dividir la escala y analizar los datos de forma agrupada, basándonos en los resultados de BSEM y en el significado de los ítems.

RESULTADOS

La versión final de la escala contó con 53 ítems agrupados en 9 factores, al realizar un análisis factorial confirmatorio multinivel (MCFA por sus siglas en inglés) el valor de χ^2 y los índices de ajuste fueron $\chi^2(547, 2578) = 4583.151$ ($p = .00$), RMSEA = .038, SRMR_{intra} = .052, SRMR_{inter} = .121, CFI = .980, and TLI = .979. Los pesos factoriales oscilaron entre .431 to .798. Con respecto a las correlaciones, a nivel intraclase variaba entre .846 (Ser Afectuoso y Cuidar a los Alumnos con Feedback) to .261 (Usar un Lenguaje No Controlador con Centrarse en el Proceso) y a nivel interclase, el valor mayor fue de 1 (Diseñar Actividades con un Nivel de Dificultad Adecuado con Promover la Utilidad y el Interés) y el menor de .755 (Usar un Lenguaje No Controlador con Promover la Utilidad y el Interés).

Para comprobar la fiabilidad de la escala utilizamos el Omega de McDonald (1999). Los datos de la Omega de McDonald deben interpretarse de forma similar al alfa de Cronbach: valores superiores a .70 - .80 son indicadores de fiabilidad. La Tabla 1 muestra que la Omega de McDonald varía entre .650 (Centrarse en el Proceso) y .893 (Feedback).

Tabla 1

Estadísticos descriptivos, correlación intraclase, fiabilidad y correlaciones.

Variable	Ítems	M	DT	ICC	ICC2	ω	1	2	3	4	5	6	7	8	9
1	9	3.788	1.468	.343	.923	.867		.958	.937	-.711	1.000	.976	.841	.927	.946
2	6	3.622	1.562	.415	.942	.818	.649		.953	-.825	.958	.886	.789	.948	.975
3	5	4.319	1.528	.361	.928	.754	.715	.658		-.824	.960	.879	.905	.932	.939
4	5	5.473	1.416	.294	.905	.739	-.386	-.371	-.366		.755	.780	.762	.886	.757
5	4	4.165	1.594	.309	.911	.747	.619	.624	.632	.384		.966	.860	.961	.948
6	4	4.740	1.444	.013	.225	.650	.280	.351	.372	.162	.324		.992	.900	.829
7	5	5.323	1.451	.180	.834	.865	.511	.415	.501	.451	.518	.262		.867	.766
8	8	4.623	1.557	.285	.901	.893	.729	.686	.740	.467	.681	.381	.610		.930
9	7	3.955	1.562	.351	.925	.867	.713	.724	.710	.425	.703	.332	.572	.813	

Nota. ICC= correlación intraclase. ω = omega de McDonald. Triángulo inferior: correlaciones intraclase. Triángulo superior: correlaciones interclase. 1=Promover la Utilidad y el Interés, 2=Identificar y Comprender los Sentimientos Negativos de los Alumnos, 3=Promover la Participación en Clase, 4=Usar un Lenguaje No Controlador, 5=Diseñar Actividades con un Nivel de Dificultad Adecuado, 6= Centrarse en el Proceso, 7=Estructurar las Clases, 8=Feedback, 9=Ser Afectuoso y Cuidar a los Alumnos.

ESTUDIO 2

MÉTODO

Participantes

En el estudio participaron 1555 alumnos de educación secundaria (51% chicas) con una edad media de 15.30 años (DT = 1.12). Los alumnos pertenecían a 82 grupos, Se trataba de alumnos que estaban entre 2º y 4º de Educación Secundaria Obligatoria, (2º de ESO, n = 588, Medad= 13.94; 3º de ESO, n = 484, Medad = 15.01; 4º de ESO, n = 483, Medad = 16.19). Al igual que en

el estudio anterior, los institutos que formaron parte del estudio son de naturaleza variada: zonas rurales, periféricas y urbanas y los estudiantes fueron, en su gran mayoría, de familias de clase media. Durante la educación secundaria los alumnos cuentan con 4 horas semanales de Matemáticas. Por último, es importante destacar que el estudio se comenzó seis meses después de que comenzara el curso por lo que los alumnos habían tenido tiempo suficiente para saber cómo era el comportamiento de su profesor en el aula.

Procedimiento

Se siguió un procedimiento similar al explicado en el estudio 1. En primer lugar, contactamos telefónicamente con los jefes de departamento de Matemáticas de varios institutos. Finalmente, nueve institutos participaron en el estudio. A la hora de realizar la evaluación, los alumnos dieron su consentimiento para formar parte del estudio y su participación fue confidencial y voluntaria. Menos del 1% rechazaron participar.

Método

Instrumentos de medida

Con el objetivo de analizar la fiabilidad de la escala utilizamos la Omega McDonald. A continuación, calculamos ICC1 para estimar qué cantidad de varianza se debía al grupo. También calculamos el ICC2 para analizar la fiabilidad de los factores como indicadores grupales. Finalmente, realizamos un MCFA para poner a prueba la estructura factorial, propuesta en el estudio anterior.

Calidad didáctica

Se utilizó la escala de 53 ítems que se describió en el estudio anterior. La fiabilidad y los valores del ICC1 se pueden consultar en la Tabla 2.

Implicación

El esfuerzo, como indicador de la implicación, fue evaluado usando cuatro ítems de la subescala *Motivated Strategies for Learning Questionnaire* (Pintrich et al., 1993), la escala estaba graduada del 1 al 7 (1 = *totalmente en desacuerdo* a 7 = *totalmente de acuerdo*). Se usaron ítems como: “Cuando el trabajo de clase es difícil, lo dejo o solo me estudio lo más fácil”. La fiabilidad y los valores del ICC1 se pueden consultar en la Tabla 2.

Notas de Matemáticas

Los centros proporcionaron las calificaciones una vez terminado el curso académico.

Análisis de datos

Se calculó la media, la desviación típica y la correlación entre las variables, tanto a nivel interclase, como a nivel intraclase. A continuación, se llevó a cabo un MSEM para obtener evidencias de la segunda hipótesis del estudio. Se probó a nivel intraclase, si la implicación predecía el rendimiento en Matemáticas y a nivel interclase, si la calidad didáctica predecía las

notas a través de la implicación. Para la imputación de los datos perdidos se utilizó el método de mínimos cuadrados ponderados por la media y la varianza (Asparouhov & Muthén, 2010).

RESULTADOS

Análisis descriptivo y correlaciones a nivel intra e interclase

Las medias, desviaciones típicas y correlaciones de todas las variables se muestran en la Tabla 2.

Tabla 2

Estadísticos descriptivos, correlación intraclase, fiabilidad y correlaciones.

Variable	M	DT	ICC1	ICC2	ω	1	2	3	4	5	6	7	8	9	10	11
1	4.033	1.683	.545	.957	.889		.946	.960	-.495	.931	.917	.874	.944	.852	.409	.232
2	3.449	1.772	.467	.942	.892	.684		.961	-.531	.966	.965	.809	.964	.911	.443	.284
3	4.436	1.655	.385	.921	.828	.618	.598		-.563	.935	.939	.852	.958	.846	.405	.181
4	2.732	1.575	.416	.930	.827	-.147	-.105	-.123		-.469	-.518	-.453	-.602	-.600	-.335	.015
5	4.123	1.736	.472	.943	.871	.532	.551	.633	-.171		.942	.847	.942	.842	.434	.209
6	4.447	1.732	.342	.905	.804	.470	.468	.544	-.140	.569		.804	.949	.846	.487	.361
7	5.181	1.668	.400	.926	.908	.519	.436	.566	-.221	.548	.571		.855	.728	.327	.121
8	4.534	1.706	.387	.921	.919	.669	.661	.721	-.224	.687	.648	.652		.923	.369	.200
9	3.907	1.663	.423	.931	.902	.660	.750	.670	-.138	.610	.526	.521	.739		.333	.242
10	4.636	1.378	.161	.782	.728	.199	.191	.220	-.119	.288	.209	.264	.285	.294		.481
11	5.217	2.212	.084	-	-	.099	.113	.113	-.091	.257	.199	.158	.183	.167	0.467	

Nota. ICC= correlación intraclase. ω = omega de McDonald. Triángulo inferior: correlaciones intra. Triángulo superior: correlaciones inter. 1=Promover la Utilidad y el Interés, 2=Identificar y Comprender los Sentimientos Negativos de los Alumnos, 3=Promover la Participación en Clase, 4=Usar un Lenguaje No Controlador, 5=Diseñar Actividades con un Nivel de Dificultad Adecuado, 6= Centrarse en el Proceso, 7=Estructurar las Clases, 8=Feedback, 9=Ser Afectuoso y Cuidar a los Alumnos, 10=Implicación, 11=Notas de Matemáticas.

Modelo multinivel

El valor de χ^2 y los índices de ajuste en el MSEM fueron $\chi^2(1504, 3170) = 23506.247$ ($p = .00$), RMSEA = .065, SRMR_{intra} = .052, SRMR_{inter} = .087, CFI = .962, and TLI = .961. En la Figura 1 se muestra que los nueve factores descritos anteriormente son indicadores de la Calidad Didáctica del profesor. A nivel intraclase, la Implicación predice las Notas de Matemáticas ($\beta = .528$; SE = .036; $p < .001$) explicando un 28% de su varianza. Por otro lado, a nivel interclase, y la Calidad Didáctica predice la Implicación ($\beta = .508$; SE = .097; $p < .001$) y esto se traduce en las Notas de Matemáticas ($\beta = .520$; SE = .171; $p < .001$); explicando un 26% y 27% de su varianza, respectivamente.

DISCUSIÓN Y CONCLUSIONES

Los profesores se enfrentan al reto de conseguir que sus alumnos aprendan y se involucren en clase. Aunque se sabe que el comportamiento del profesor durante las lecciones se relaciona

con el funcionamiento académico de los alumnos; desde la teoría de la autodeterminación no se han precisado cuáles son los comportamientos específicos de los profesores que optimizan el rendimiento de los alumnos (Stroet et al., 2013, 2015). Por tanto, en el primer estudio, se diseñó un instrumento para que los alumnos puedan evaluar los comportamientos específicos de los profesores y así, evaluar su calidad didáctica. En el segundo estudio, se observó que la calidad didáctica de los profesores predecía el rendimiento de los alumnos en Matemáticas.

Concretamente, en el estudio 1, se desarrolló una escala que evalúa la calidad didáctica del profesor en base a nueve factores relacionados con los comportamientos de los docentes. El uso de la escala ha proporcionado evidencias en cuanto a su fiabilidad y validez. Sin embargo, se necesita evidenciar que la escala tiene una estructura válida probándola con otra muestra y así demostrar que predice la implicación y rendimiento de los alumnos, para lo que se plantea el estudio 2.

El estudio 2 se buscó aportar evidencias de validez predictiva de la escala previamente diseñada. Se observó que la calidad didáctica predecía el rendimiento académico de los alumnos, a través del esfuerzo (como indicador de la implicación). Es decir, la implicación se comportaba como una variable mediadora entre el comportamiento del profesor (la calidad didáctica) y el rendimiento en Matemáticas.

Para finalizar, nos gustaría destacar que la escala desarrollada en esta investigación puede ser un instrumento útil para evaluar la calidad didáctica de una forma fiable y válida. Además, esta escala puede servir para como punto de partida para desarrollar intervenciones con el objetivo de mejorar la implicación y el rendimiento de alumnos de secundaria.

REFERENCIAS BIBLIOGRÁFICAS

- Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeier et al. *Journal of Management*, 41(6), 1561–1577. <http://doi.org/10.1177/0149206315591075>
- Asparouhov, T., & Muthén, B. O. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397–438. <http://doi.org/10.1080/10705510903008204>
- Asparouhov, T., & Muthén, B. O. (2010). *Weighted least squares estimation with missing data*. Retrieved from <http://www.statmodel.com/download/GstrucMissingRevision.pdf>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, US: Plenum.
- Deci, E. L., & Ryan, R. M. (2000). The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268. http://doi.org/10.1207/S15327965PLI1104_01

- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185.
- Kiemer, K., Gröschner, A., Pehmer, A.-K., & Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction*, 35, 94–103. <http://doi.org/10.1016/j.learninstruc.2014.10.003>
- Liu, W. C., Wang, J. C. K., & Ryan, R. M. (2016). *Building autonomous learners*. Singapore: Springer.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- Muthén, B. O., & Asparouhov, T. (2012). Bayesian SEM: A more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313–335. <http://doi.org/10.1037/a0026802>
- Núñez, J. L., & León, J. (2015). Autonomy support in the classroom: A review from the self-determination theory. *European Psychologist*, 20(4), 275–283.
- Pintrich, P. R., Smith, D. A. F., Garcia, T., McKeachie, W. J., García, T., & McKeachie, W. J. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3), 801–813. <http://doi.org/10.1177/0013164493053003024>
- Stroet, K., Opdenakker, M. C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review*, 9, 65–87. <http://doi.org/10.1016/j.edurev.2012.11.003>
- Stroet, K., Opdenakker, M. C., & Minnaert, A. (2015). What motivates early adolescents for school? A longitudinal analysis of associations between observed teaching and motivation. *Contemporary Educational Psychology*, 42, 129–140. <http://doi.org/10.1016/j.cedpsych.2015.06.002>