

Perfiles docentes asociados a las prácticas de evaluación en matemáticas

Teaching Profiles Associated with Assessment Practices in Mathematics

Angélica Fabiana Oviedo Mandujano*¹, Coral González Barbera** y Joaquín Caso Niebla***

*Facultad de Ciencias Humanas. Universidad Autónoma de Baja California (México)

**Facultad de Educación. Universidad Complutense de Madrid (España)

***Instituto de Investigación y Desarrollo Educativo. Universidad Autónoma de Baja California (México)

Resumen

Esta investigación tiene como objetivo identificar los perfiles docentes asociados a las prácticas de evaluación en matemáticas y su diferenciación en términos de variables sociodemográficas. Se contó con una muestra de 4,664 docentes ($M_{edad}=38$, $DE=9.70$, 77.36% mujeres) de educación primaria en la entidad de Baja California, México. Para el análisis de datos se realizó un Análisis de Clases Latentes (ACL) y se utilizaron pruebas t-student y chi-cuadrada. Se identificaron dos perfiles docentes, el intensivo y el moderado, cuya distinción radica en la frecuencia en la que se realizan las distintas prácticas de evaluación documentadas. Si bien, ambos perfiles se encuentran claramente diferenciados en la mayoría de las dimensiones del proceso evaluativo, no fue así para las variables sobre el uso de técnicas e instrumentos de evaluación. Además, se ubicaron diferencias en las variables sociodemográficas de sexo, edad, años de experiencia docente, institución de formación inicial, participación en espacios de formación continua, sostenimiento de la escuela y asentamiento rural o urbano del centro escolar, aunque con tamaño del efecto bajos. No obstante, el grado de estudios, grado escolar, tipo de jornada y la modalidad de la escuela no representaron diferencias entre perfiles. Se acepta parcialmente la hipótesis en lo relativo a que las variables sociodemográficas permiten caracterizar perfiles docentes asociados

1 **Correspondencia:** Angélica Fabiana Oviedo Mandujano, angelica.oviedo@uabc.edu.mx, Facultad de Ciencias Humanas, Universidad Autónoma de Baja California, 21100, Mexicali, Baja California, México.

a las prácticas de evaluación en matemáticas donde, aunque su efecto es bajo, pueden orientar la implementación de estrategias diferenciadas para apoyar la labor docente.

Palabras clave: evaluación; práctica docente; perfiles; educación matemática.

Abstract

This research aims to identify teaching profiles associated with assessment practices in Mathematics and their differentiation in terms of sociodemographic variables. A sample of 4,664 teachers ($M_{age}=38$, $SD=9.70$, 77.36% women) at the primary educational level in the state of Baja California, Mexico was used. For data analysis, a Latent Class Analysis (LCA) was performed, and t-student and chi-square tests were conducted. Two teaching profiles were identified, the intensive and the moderate, whose distinction lies in the frequency with which they carry out the different documented evaluation practices. Although both profiles are clearly differentiated in most dimensions of the evaluation process, this was not the case for the variables on the use of evaluation techniques and instruments. In addition, differences were found in the sociodemographic variables of sex, age, years of teaching experience, institution of initial training, participation in continuous training courses, school funding and rural or urban location of the school, although with low effect sizes. However, the education level, school grade, type of working day and type of school did not represent differences between profiles. The hypothesis is partially accepted, which states that sociodemographic variables allow characterizing teaching profiles associated with assessment practices in Mathematics where, although their effect is low, they can guide the implementation of differentiated strategies to support the teaching work.

Keywords: assessment; teaching practices; profiles; Mathematics education.

Introducción

Las prácticas docentes de evaluación en matemáticas

Las prácticas docentes de evaluación han adquirido relevancia en la investigación educativa, dado el reconocimiento de su capacidad para condicionar la enseñanza e incidir en el aprendizaje (Navarro-Mosquera et al., 2017, Tariq et al., 2023). Además, en las últimas décadas se han posicionado con mayor fuerza dentro de las políticas educativas en las que se aboga por un enfoque de corte constructivista (Acar-Erdol y Yildizli, 2018; Schildkamp et al., 2020).

Cuando la evaluación se realiza en el marco del aprendizaje de las matemáticas, las prácticas evaluativas se entienden como las acciones que despliega el profesorado para obtener información del aprendizaje de sus estudiantes respecto al nivel de conocimiento matemático y su competencia para resolver problemas. Esto requiere de un proceso sistemático orientado por objetivos y criterios definidos durante la planeación, a fin de realizar interpretaciones que les permita tomar decisiones y la comunicación con otros actores.

Dentro de la evaluación es posible identificar al menos cinco dimensiones con base en las cuales estructurar este proceso: finalidad de la evaluación, objeto de evaluación,

estrategias de evaluación, interpretación de evidencias de aprendizaje, y comunicación y uso de resultados (Chappuis et al., 2012; Demosthenous et al., 2021; Leenknecht et al., 2021; Sheikh y Manap, 2024).

En relación con la finalidad de la evaluación, suele dividirse en dos grandes vertientes referidas comúnmente con los términos acuñados por Scriven en 1967 relativos a la evaluación sumativa y formativa. La primera hace alusión a la función social de la evaluación orientada a la certificación de aprendizajes y, la segunda, a una función de índole pedagógico (Ravela et al., 2017; Schellekens et al., 2021), asociada a la mejora de la enseñanza y el aprendizaje, que promueve el desarrollo de la metacognición y la capacidad de autorregulación (Braund y DeLuca, 2018; Zulliger et al., 2022).

Para el caso de la evaluación formativa, Black y Wiliam (2009) propusieron cinco estrategias clave que permiten desarrollar una evaluación de esta naturaleza: compartir y clarificar los aprendizajes esperados y los criterios de evaluación; diseñar e implementar actividades que ofrezcan evidencia del aprendizaje obtenido; proporcionar retroalimentación; activar a los y las estudiantes como fuente de aprendizaje para sus pares y; activar a cada estudiante como responsable de su propio aprendizaje. Estas estrategias continúan siendo un referente en estudios más recientes como los de Chan (2021) y Wafubwa y Csikos (2021).

Por otra parte, la evaluación siempre está referida a un conjunto de contenidos o habilidades específicos que constituyen su objeto. En el caso de las matemáticas, el enfoque de resolución de problemas se ha posicionado como el marco ideal para su enseñanza, desde el que se promueve el desarrollo de operaciones mentales como el análisis, síntesis, generalización, abstracción, y el desarrollo del pensamiento heurístico, flexible y creativo (Herbert, 2021). Los planteamientos de este enfoque encuentran sus inicios desde las aportaciones de Polya (1965) quien estructuró el proceso de resolución de problemas en las etapas de comprensión del problema, planeación de la solución, ejecución del plan y evaluación; las cuales son susceptibles de formar parte el objeto a evaluar en correspondencia con el enfoque didáctico.

En cuanto a la evaluación sumativa, que también forma parte de las tareas fundamentales del profesorado, existen diversos mecanismos para realizar la interpretación de las evidencias de aprendizaje a fin de generar una nota. Un aspecto de interés tiene que ver con la combinación de aspectos de índole cognitivo y no cognitivo. Al respecto, en la literatura especializada se señala que el considerar elementos no académicos dentro de la calificación reduce su validez (Dagdag y Dagdag, 2020), por lo que esta práctica no es recomendable para fortalecer el aprendizaje desde la evaluación (Chappuis et al., 2012; Paderes y Ramos, 2024).

Finalmente, la dimensión de comunicación de los resultados de la evaluación se refiere a las formas en que el profesorado informa a diferentes actores del aprendizaje obtenido (Dagdag y Dagdag, 2020), entre los que se encuentran el estudiantado, el colectivo escolar, y en el caso de la educación primaria, figuran de manera importante las familias, dado su papel clave en el desempeño del alumnado (Lara y Saracostti, 2019). En correspondencia con la finalidad de la evaluación, los resultados se podrán comunicar con fines meramente informativos o acompañados de estrategias de mejora para la enseñanza y el aprendizaje.

Planteamiento del problema

En el caso de México, la resolución de problemas en matemáticas y el enfoque de la evaluación formativa se incorporaron al discurso pedagógico de los planes de estudio de primaria desde la reforma realizada en 1993. Sin embargo, de forma histórica, los resultados de evaluaciones a gran escala dan cuenta de un bajo logro de los aprendizajes en esta área del conocimiento.

Muestra de ello, son los resultados educativos en Matemáticas registrados en la prueba PLANEA en 2018, en los que cerca de la mitad del estudiantado de sexto grado de primaria (49.1%) se ubicó en un nivel de conocimiento insuficiente, lo que representa serias dificultades para continuar con su trayectoria escolar (Instituto Nacional para la Evaluación de la Educación [INEE], 2018). Estos resultados son consistentes con los reportados recientemente por el Programa para la Evaluación Internacional de los Estudiantes (PISA) de la Organización para la Cooperación y el Desarrollo Económico (OCDE), donde el 65.8% de estudiantes de 15 años en México se ubicó por debajo del nivel mínimo en Matemáticas, además de registrar un retroceso respecto a la evaluación de 2018 (OCDE, 2023).

Ante esta situación, y dada la relevancia de la evaluación en la adquisición de aprendizajes, resulta de interés indagar sobre las prácticas docentes asociadas a este proceso. Una forma de abordarlo es el análisis de perfiles que permita identificar patrones en el profesorado (Sayac y Veldhuis, 2021; Veldhuis y Van Den Heuvel-Panhuizen, 2014; Zhao et al., 2018). En este marco, se plantean como preguntas de investigación: ¿qué perfiles docentes se identifican en las prácticas de evaluación en matemáticas? y, ¿qué variables permiten diferenciar estos perfiles? Lo anterior deriva en el objetivo de identificar perfiles que caractericen las prácticas de evaluación en matemáticas en educación primaria, y su diferenciación en términos de variables sociodemográficas, planteando como hipótesis que estas últimas representan distinciones entre grupos de docentes.

Método

Participantes

El estudio se realizó en Baja California, México, con 14,276 docentes de primaria. Mediante un muestreo por conveniencia, se obtuvo una participación de 5,665 profesores(as), que representaron al 39.68% de la población. Se eliminaron, mediante el método *listwise*, los casos con valores perdidos, restando finalmente 4,674 participantes de los cuales 77.36% son mujeres, con una edad media de 38 años (DE=9.70), y presentan estudios principalmente de licenciatura (81.64%) y más de 6 años de experiencia en este nivel educativo (65.25%).

Instrumento

Se utilizó un cuestionario de autoinforme conformado por 11 ítems con información de carácter sociodemográfico y una escala sobre prácticas de evaluación en matemáticas, de la cual se consideraron 47 ítems con opción de respuesta tipo Likert que exploran la frecuencia con la que el profesorado realiza cada una de las afirmaciones planteadas.

Para la escala se obtuvieron evidencias de validez de contenido, mediante juicio de expertos, que permitió realizar ajustes al instrumento en su etapa inicial. Posteriormente, se obtuvieron evidencias relativas a la estructura interna, para lo cual se calculó el coeficiente de alfa ordinal, presentando un valor de .97, considerado como un nivel excelente de consistencia interna (George y Mallery, 2003), además de la correlación punto biserial corregida, con valores mayores a .30, que indican una buena capacidad de discriminación de los ítems (Ebel y Frisbie, 1986). Asimismo, se empleó el análisis factorial exploratorio (AFE) y confirmatorio (AFC) tanto con fines de validación como de reducción de datos.

Para el AFC, se introdujeron al modelo aquellos ítems que pertenecían a una dimensión teórica con una clasificación *per se*. En este procedimiento se utilizó el método de estimación de mínimos cuadrados no ponderados sobre la matriz de correlación policórica y se revisaron los valores del índice de bondad de ajuste (GFI $\geq .95$), la raíz cuadrada de la media del error de aproximación (RMSEA $\leq .06$), la raíz cuadrada media estandarizada (SRMR $\leq .08$), el índice de ajuste comparativo (CFI $\geq .96$) y el índice de Tucker Lewis (TLI $\geq .95$), siguiendo a Lloret-Segura et al. (2014)

Por su parte, el AFE se utilizó con aquellos ítems que no presentaban una clasificación previa, para lo cual se empleó el método de extracción de mínimos cuadrados no ponderados y rotación varimax, posterior a la obtención de valores aceptables en la prueba de esfericidad de Bartlett $p < 0.001$ (Montoya, 2007) y en el estadístico Kaiser-Meyer-Olkin $KMO > 0.80$ (Lloret-Segura et al., 2014). Se revisó que las comunalidades fueran superiores a .30, y se mantuvieron cargas factoriales por encima de .30, además de corroborar que los factores obtenidos tuvieran una explicación teórica coherente.

Derivado de lo anterior, se obtuvieron 11 factores de este procedimiento, cuyos valores resultantes permiten corroborar que la estructura teórica presenta una alta correspondencia con los datos (ver Tabla 1).

Tabla 1

Variables resultantes del análisis factorial

Dimensión teórica	Número de ítems	Factor	Análisis factorial
Finalidad de la evaluación	4	Finalidad formativa	AFC
	4	Finalidad sumativa	
Objeto de evaluación	4	Comprensión del problema y capacidad comunicativa	AFE
	3	Ajuste y evaluación de la solución del problema	
Estrategias de evaluación formativa	6	Estrategias de evaluación formativa	AFC
	8	Técnicas e instrumentos de evaluación alternativos	AFE
	3	Exámenes como instrumentos de evaluación	

Dimensión teórica	Número de ítems	Factor	Análisis factorial
Interpretación de las evidencias	3	Aspectos cognitivos que componen la calificación	AFC
	5	Aspectos no cognitivos que componen la calificación	
Comunicación y uso de resultados	4	Comunicación con el colectivo escolar	AFE
	3	Comunicación con estudiantes y familias	
47			

Nota. En el AFC se introdujeron los ítems de acuerdo con las dimensiones teóricas propuestas en un modelo de rasgos correlacionados obteniendo índices de bondad de ajuste óptimos ($\chi^2=3,495.059$, $gl=199$, $GFI=.996$, $CFI=.981$, $TLI=.978$, $RMSEA=.060$, $90\% CI [.058-.061]$, $Valor-p RMSEA<.001$, $SRMR=.054$). En el caso del AFE, la proporción de varianza explicada de las tres soluciones factoriales fue de 71%, 49% y 60%, respectivamente.

Procedimiento

Se realizó una aplicación censal mediante una plataforma digital desarrollada para este fin. La invitación al profesorado se realizó vía correo electrónico a través de la Secretaría de Educación de Baja California y estuvo abierta del 3 al 27 de mayo de 2022. El cuestionario incluyó una sección inicial dedicada a explicar la confidencialidad y anonimato de la información recuperada. En este espacio se solicitó a cada docente otorgara su consentimiento para el uso de datos, citando la normativa que regula el tratamiento de los datos personales en México.

Análisis de datos

Para responder a la pregunta de investigación relativa a ¿qué perfiles docentes se identifican en las prácticas de evaluación en matemáticas?, se utilizó el ACL, técnica estadística perteneciente a la familia de los modelos mixtos que permite la identificación de clases subyacentes de individuos en función de las diferencias en sus respuestas a los ítems de una prueba (Nylund-Gibson y Choi, 2018; Ondé y Alvarado, 2019). En este procedimiento los sujetos se agrupan por asignación modal, es decir, en la clase donde tienen la probabilidad más alta de pertenecer de acuerdo con sus patrones de respuesta (Bauer, 2022).

Para lo anterior, se utilizó el paquete *poLCA* (Linzer y Lewis, 2011) del software R, que utiliza el algoritmo de Expectativa—Maximización (EM) para maximizar la función de probabilidad logarítmica del modelo. Dado que los resultados varían debido a la inicialización aleatoria de este algoritmo y a fin de contrarrestar el riesgo de identificar un máximo local en lugar de uno global en la función de verosimilitud, se estimó cada modelo con al menos 50 conjuntos de valores iniciales aleatorios y se fijó el número de iteraciones en 3,000, que aseguraron la convergencia en soluciones similares y por lo tanto la estabilidad de la clasificación.

Para determinar el número de perfiles se consideraron tanto criterios estadísticos como teóricos. En relación con los criterios estadísticos se observaron los valores más bajos en tres índices de ajuste: criterio de información bayesiano (BIC), su variante ajustada al tamaño de la muestra (SABIC) y el criterio de información de Akaike (AIC), de acuerdo con lo señalado por Weller et al. (2020). Del modelo seleccionado, se calculó la entropía y las probabilidades posteriores promedio de cada clase para corroborar que se obtuvieran valores adecuados.

Posteriormente, a fin de responder a la pregunta ¿qué variables permiten diferenciar los perfiles docentes asociados a las prácticas de evaluación?, se utilizó la prueba t-student —junto al cálculo de la d de Cohen— para observar las diferencias en cuanto a los 11 factores sobre las prácticas docentes de evaluación y la variable de edad. Además, se empleó la prueba chi-cuadrado para el resto de las variables sociodemográficas de naturaleza ordinal o nominal. Estos análisis se realizaron en el software SPSS 27, y el cálculo del tamaño del efecto de la prueba chi-cuadrado a través del programa G*Power.

Resultados

Respecto a la identificación de los perfiles docentes asociados a las prácticas de evaluación se seleccionó el modelo de dos clases latentes, ya que, siguiendo a Bauer (2022) y Nylund-Gibson y Choi (2018), es donde se localiza el punto inflexión que muestra la disminución más prominente de los valores en los criterios BIC, SABIC y AIC, registrando una disminución con valores marginales para cada clase adicional (Figura 1).

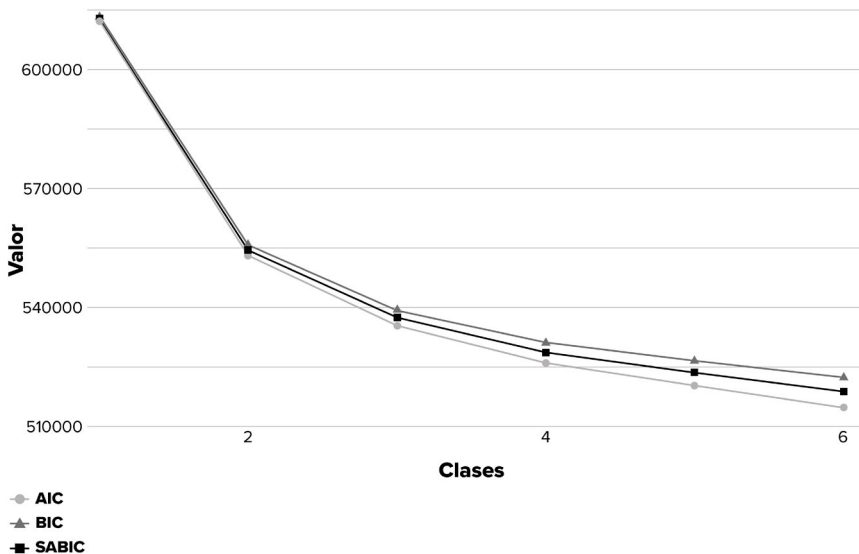


Figura 1. Diagrama de codo con los valores del BIC, SABIC y AIC para modelos de una a seis clases latentes

Nota. BIC=criterio de información bayesiano; SABIC=criterio de información bayesiano ajustado al tamaño de la muestra y AIC=criterio de información de Akaike.

Para el modelo de dos clases, se calculó la entropía y las probabilidades posteriores promedio de cada clase. En relación con la entropía, se esperan valores mínimos de .80 (Bauer, 2022; Weller et al., 2020), obteniendo un resultado de .90, lo que indica que las dos clases latentes se encuentran claramente separadas. Por su parte, las probabilidades posteriores promedio de cada clase, resultaron de .99, lo cual aportó evidencia adicional para optar por esta clasificación, ya que de acuerdo con Masyn (2013) son deseables valores mayores a .70.

Las dos clases latentes obtenidas, denominadas *perfiles docentes asociados a las prácticas de evaluación en matemáticas*, dividieron a la muestra en un 55% de docentes pertenecientes al perfil 1 (2,558) y un 45% al perfil 2 (2,116). Del análisis de las probabilidades de respuesta condicional, se observó que el profesorado que integra el perfil 1 generalmente presenta probabilidades más altas de responder a las categorías de *siempre* y *casi siempre* que el del perfil 2.

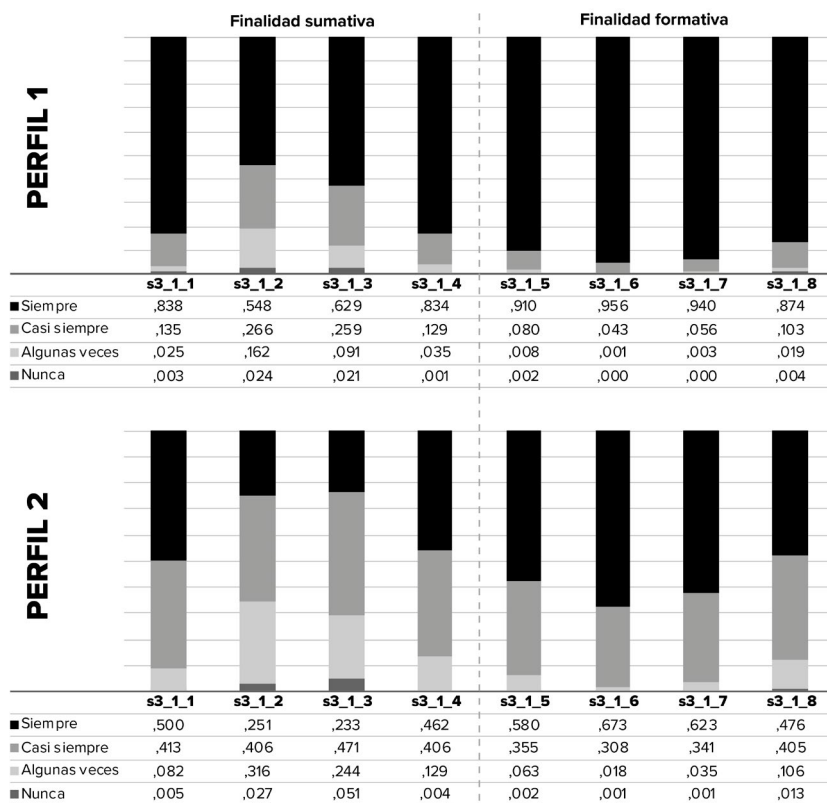


Figura 2. Probabilidades condicionales de los ítems asociados a la finalidad de la evaluación
 Nota. S3_1_1=Medir su aprendizaje; S3_1_2=Asignar una calificación; S3_1_3= Nivelar al grupo; S3_1_4=Comunicar a las familias el avance de sus hijos; S3_1_5=Identificar sus conocimientos previos; S3_1_6=Orientarlos en su aprendizaje; S3_1_7=Adaptar mis estrategias de enseñanza; S3_1_8=Identificar mis propias necesidades de actualización docente.

En la Figura 2 se da cuenta de estos resultados, donde a partir de los ítems asociados a la finalidad de la evaluación, se identifica que el profesorado del perfil 1 tienen una probabilidad superior al 80% de responder en la categoría *siempre*, excepto en el caso de la evaluación para asignar una calificación y nivelar al grupo con un 55% y 63%, respectivamente, donde la categoría *siempre* sigue concentrando la probabilidad más alta. Por su parte, en el perfil 2, las probabilidades se distribuyeron entre las categorías *siempre* y *casi siempre*, y para los ítems relativos a la evaluación para asignar calificación y nivelar el grupo, la categoría *algunas veces* también obtuvo valores de probabilidades destacables.

Situación similar se observó en los ítems de los factores asociados al objeto de evaluación (Figura 3), el factor de estrategias de evaluación formativa (Figura 4), los relacionados con la asignación de calificación (Figura 5), así como los relativos a la comunicación de resultados (Figura 6).

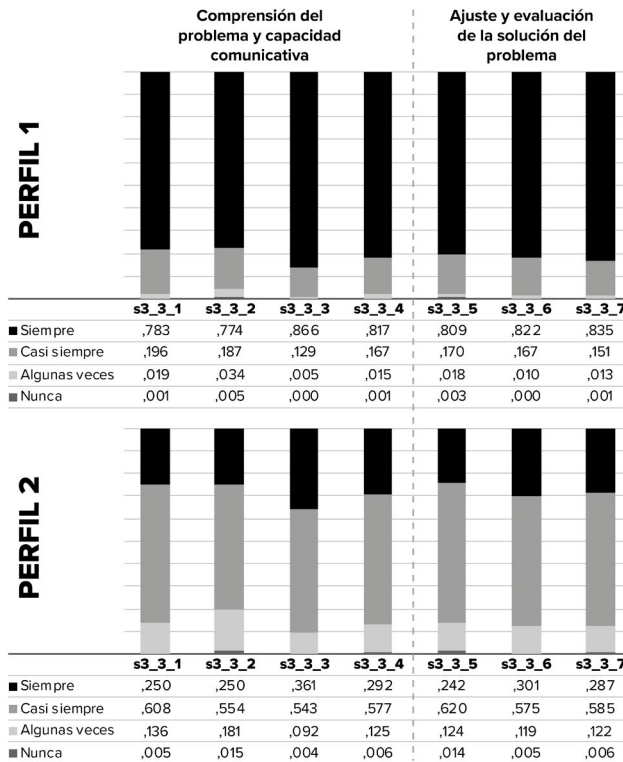


Figura 3. Probabilidades condicionales de los ítems asociados al objeto de evaluación

Nota. S3_3_1=Identifiquen las relaciones entre los datos explícitos e implícitos proporcionados en el planteamiento de un problema; S3_3_2=Expliquen de qué trata un problema con sus propias palabras; S3_3_3=Visualicen qué pasos van a seguir antes de resolver un problema; S3_3_4=Sean capaces de explicar cada paso que realizan al resolver un problema; S3_3_5=Ajusten o modifiquen los pasos para resolver un problema cuando se les presenta alguna dificultad; S3_3_6=Comprueben que los resultados obtenidos efectivamente resuelven el problema planteado; S3_3_7=Identifiquen otros modos de resolver un mismo problema.

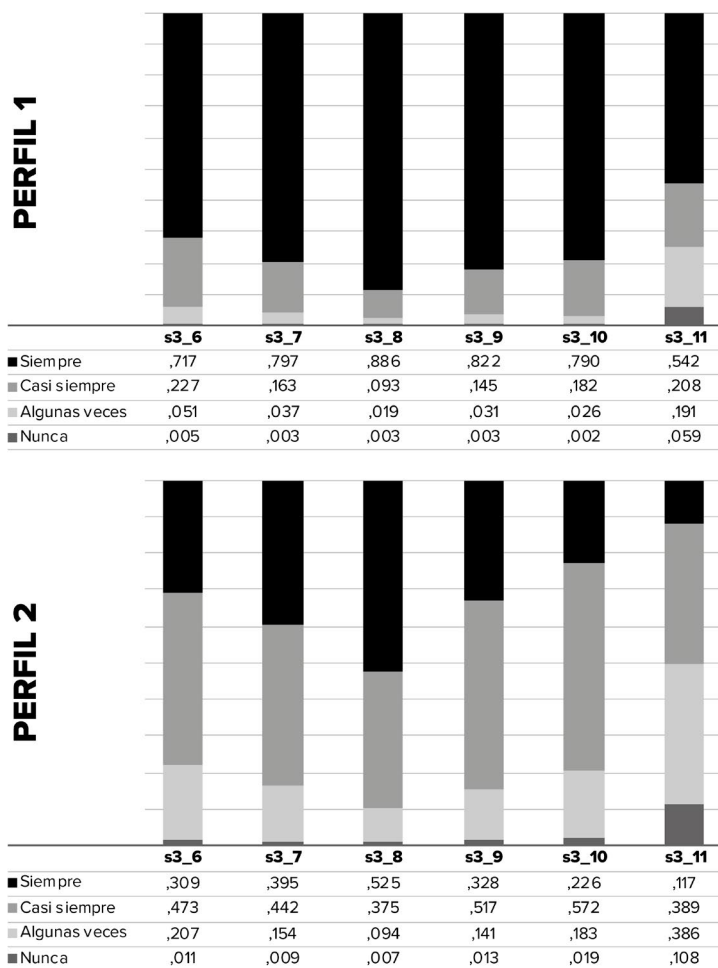


Figura 4. Probabilidades condicionales para los ítems asociados a estrategias de evaluación formativa

Nota. S3_6=Comunico a mis estudiantes qué se espera que aprendan en cada bloque temático; S3_7=Explico a mis estudiantes cómo van a ser evaluados; S3_8=Informo a las familias cómo serán evaluados sus hijos; S3_9=Solicito a mis estudiantes que revisen sus trabajos o tareas para que identifiquen sus aciertos y errores, antes de proporcionarles retroalimentación; S3_10=Solicito a mis estudiantes que reflexionen sobre sus fortalezas y oportunidades de aprendizaje; S3_11=Solicito a mis estudiantes que revisen los trabajos o tareas de sus compañeros para que identifiquen las fortalezas y áreas de oportunidad.

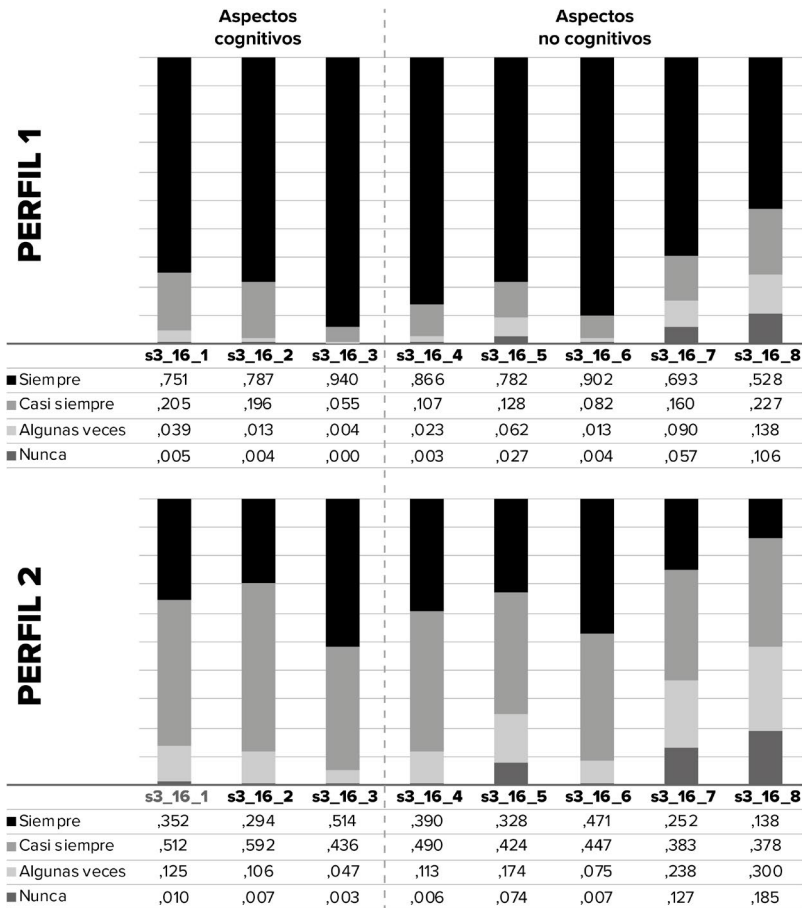


Figura 5. Probabilidades condicionales para los ítems asociados a aspectos que componen la calificación

Nota. S3_16_1=Las respuestas correctas en los exámenes, actividades en clase o tareas.; S3_16_2=El proceso de solución seguido en los exámenes, actividades en clase o tareas.; S3_16_3=La participación en clase.; S3_16_4=El desempeño en los trabajos en equipo.; S3_16_5=La asistencia a clase.; S3_16_6=La entrega de trabajos y tareas en tiempo y forma.; S3_16_7=La disciplina/conducta en el aula.; S3_16_8=El interés y gusto por las matemáticas.

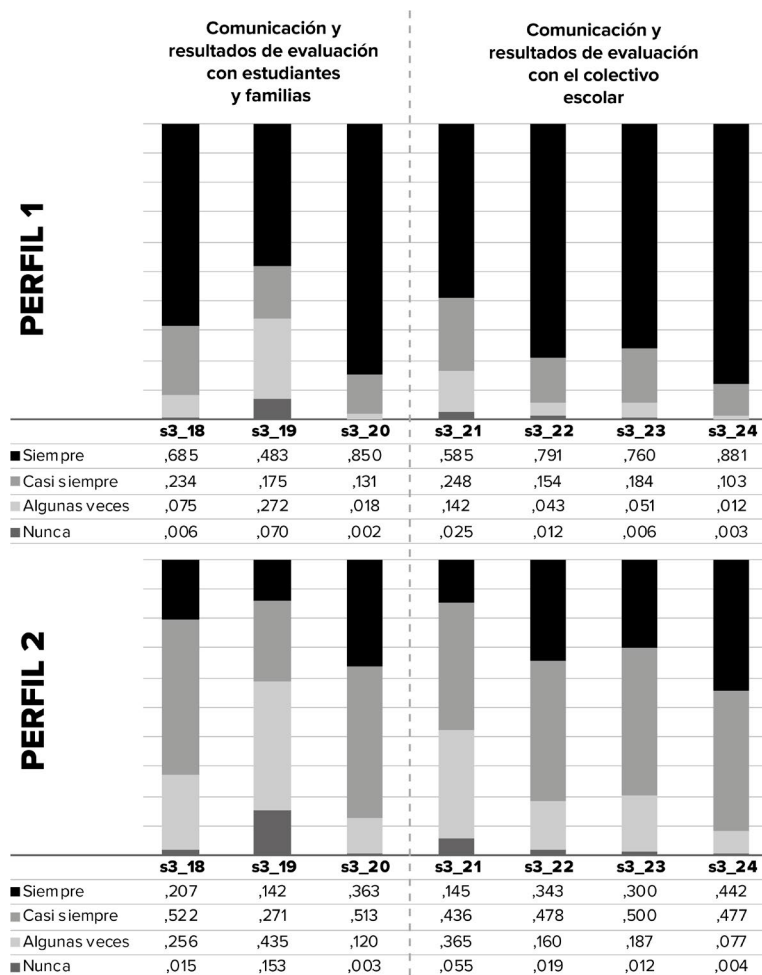


Figura 6. Probabilidades condicionales para los ítems asociados a la comunicación de resultados
 Nota. S3_18=Dedico tiempo para dialogar con mis estudiantes acerca de su calificación; S3_19=Entrego un reporte escrito (diferente a la boleta) a las familias; S3_20=Informo a las familias cómo pueden ayudar a sus hijos a mejorar su aprendizaje; S3_21=Obtengo retroalimentación de mis colegas docentes para mejorar el aprendizaje de mis estudiantes; S3_22=Comparto los resultados de evaluación con el(la) docente que atenderá mi grupo en el siguiente grado escolar; S3_23=Comparto los resultados de evaluación en las reuniones del colectivo escolar; S3_24=Los resultados de aprendizaje de mis estudiantes son tomados en cuenta en la planeación escolar.

No obstante, en el caso de los ítems que conforman los factores de *técnicas e instrumentos de evaluación alternativos* y *exámenes como instrumentos de evaluación*, se observa que las probabilidades más altas no se presentan para todos los casos en las mismas categorías (Figura 7). Por ejemplo, en el caso de los exámenes escritos, el profesorado del perfil 1 tiene una probabilidad superior al 40% de responder que *siempre*

aplica exámenes escritos tanto de respuesta abierta como cerrada, mientras que el profesorado del perfil 2 tendrá mayores probabilidades de responder que *casi siempre* aplican exámenes de preguntas abiertas, pero solo *algunas veces* los de respuesta cerrada.

Para los instrumentos agrupados bajo el nombre de alternativos, aun cuando el perfil 1 presenta probabilidades más altas en la categoría *siempre* que el perfil 2, se observan al interior de cada ítem probabilidades similares entre categorías, como es el caso de los debates o discusiones, donde existen probabilidades entre 27.1% a 35% en *algunas veces*, *casi siempre* o *siempre*. Por su parte, tendencias observadas en el perfil 2 varían, donde el uso de portafolios de evidencias, debates o discusiones, guías de observación y rúbricas se ubican principalmente en las categorías *algunas veces* y *casi siempre*, mientras que las preguntas durante la clase se aplican con mayor probabilidad *casi siempre* y *siempre*. Los instrumentos en el que coinciden ambos perfiles son el cuaderno, libro de texto, libros complementarios o guías de apoyo, donde en la categoría *siempre* se presentan las más altas probabilidades de respuesta.

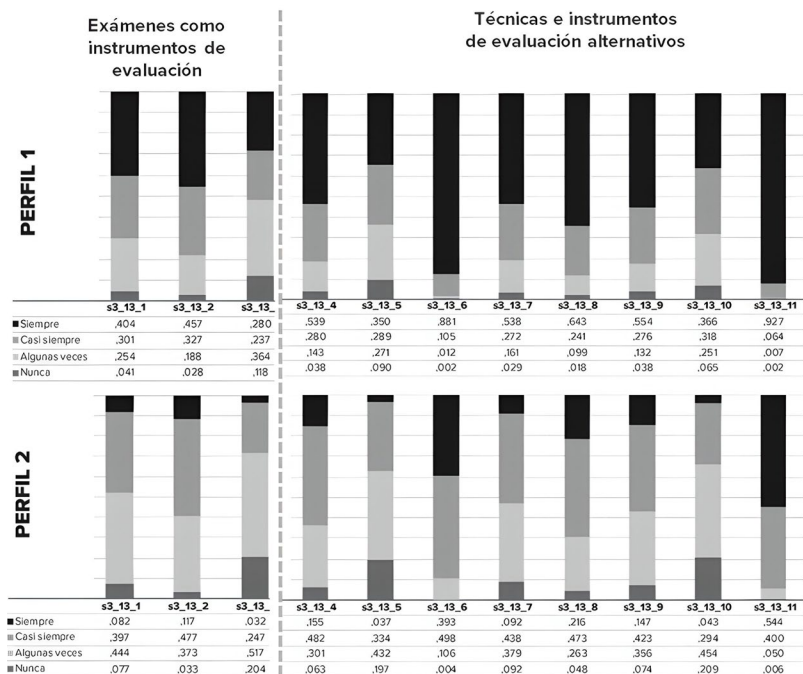


Figura 7. Probabilidades condicionales de los ítems asociados a técnicas e instrumentos de evaluación

Nota. S3_13_1=Exámenes escritos con preguntas de respuesta abierta; S3_13_2=Exámenes escritos con preguntas de respuesta cerrada o múltiple; S3_13_3=Exámenes orales. S3_13_4=Portafolios de evidencias; S3_13_5=Debates o discusiones en torno a conceptos y procedimientos matemáticos; S3_13_6=Preguntas durante la clase; S3_13_7=Guías de observación; S3_13_8=Listas de cotejo; S3_13_9=Rúbricas; S3_13_10=Escalas de actitudes; S3_13_11=Cuaderno, libro de texto, libros complementarios o guías de apoyo.

Los resultados presentados, se corroboran al realizar la prueba t-student donde el perfil 1 puntuó por encima de la media en todos los factores, mientras que el perfil 2 se ubicó por debajo de ésta. No obstante, las pruebas reportan diferencias significativas para todos los casos, con excepción de los factores correspondientes a la aplicación de instrumentos de evaluación, ya sean alternativos o de tipo examen. Las diferencias encontradas presentan tamaños del efecto grandes, oscilando entre .884 y 1.724, de acuerdo con Cárdenas y Arancibia (2014) (ver Tabla 2 y Figura 8).

A partir del comportamiento observado en los dos perfiles docentes se destaca que sus diferencias estriban en la frecuencia más que en el tipo de prácticas evaluativas. En este sentido, se optó por agregar la etiqueta de *intensivo* al perfil 1 y *moderado* al perfil 2, ya que estos últimos, si bien registran una frecuencia menor, ubican gran parte de sus respuestas en la categoría *casi siempre*.

Tabla 2

Resultados de la prueba t-student entre las variables de prácticas de evaluación

Factores	Perfil 1 Intensivo (n= 2558)			Perfil 2 Moderado (n= 2116)					
	M	DE	Media de error estándar	M	DE	Media de error estándar			
p1	0.55	0.81	0.02	-0.67	0.78	0.02			
p2	0.57	0.72	0.01	-0.69	0.84	0.02			
p3	0.42	0.68	0.01	-0.51	1.09	0.02			
p4	0.43	0.72	0.01	-0.51	1.05	0.02			
p5	0.59	0.76	0.01	-0.72	0.76	0.02			
p6	0.02	1.00	0.02	-0.02	1.00	0.02			
p7	0.01	1.00	0.02	-0.01	1.00	0.02			
p8	0.60	0.73	0.01	-0.73	0.77	0.02			
p9	0.56	0.79	0.02	-0.68	0.78	0.02			
p10	0.37	0.79	0.02	-0.44	1.05	0.02			
p11	0.39	0.87	0.02	-0.47	0.94	0.02			
Factores	t	gl	p	Diferencia de medias	Diferencia de error estándar	d	LI	LS	
p1	52.71	4,562.59	>.001	1.22	0.02	1.54	1.48	1.61	
p2	53.93	4,186.01	>.001	1.26	0.02	1.61	1.54	1.67	
p3	34.41	3,400.03	>.001	0.93	0.03	1.05	0.99	1.12	

Factores	t	gl	p	Diferencia de medias	Diferencia de error estándar	d	LI	LS
p4	34.96	3,604.36	>.001	0.94	0.03	1.06	1.00	1.13
p5	58.64	4,497.60	>.001	1.31	0.02	1.72	1.66	1.79
p6	1.42	4,672.00	0.156	0.04	0.03	0.04	-0.02	0.10
p7	0.53	4,672.00	0.597	0.02	0.03	0.02	-0.04	0.07
p8	59.94	4,415.13	>.001	1.33	0.02	1.77	1.70	1.84
p9	53.92	4,543.90	>.001	1.24	0.02	1.58	1.52	1.65
p10	29.30	3,840.35	>.001	0.81	0.03	0.88	0.82	0.94
p11	32.54	4,672.00	>.001	0.86	0.03	0.96	0.90	1.02

Nota. N=4,674. p1=Finalidad sumativa; p2=Finalidad formativa; p3=Comprensión del problema y capacidad comunicativa; p4=Ajuste y evaluación de la solución del problema; p5=Estrategias de evaluación formativa; p6=Exámenes como instrumentos de evaluación; p7=Técnicas e instrumentos de evaluación alternativos; p8=Aspectos cognitivos que componen la calificación; p9=Aspectos no cognitivos que componen la calificación; p10= Comunicación con el colectivo escolar; p11= Comunicación con estudiantes y familias; gl=grados de libertad; d= d de Cohen; LI=Límite inferior del tamaño del efecto; LS=Límite superior del tamaño del efecto.

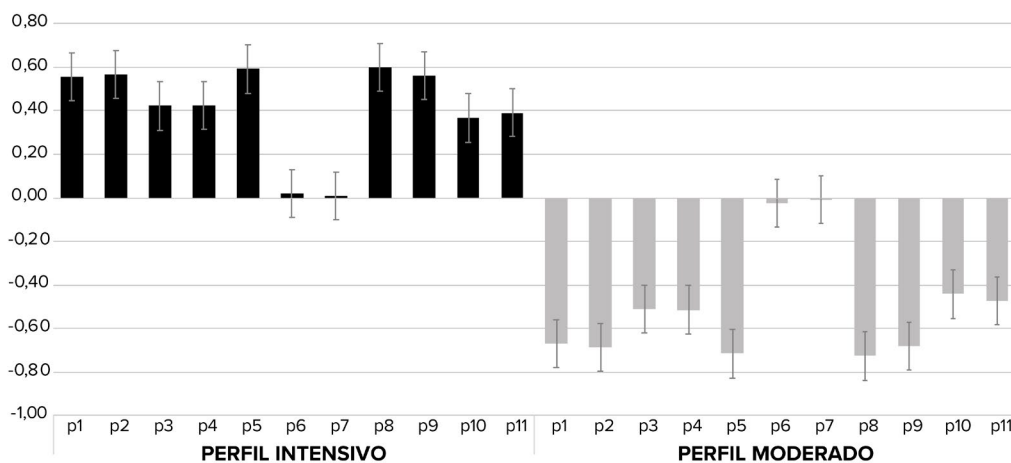


Figura 8. Comportamiento de las variables de prácticas de evaluación

Nota. N=4,674. p1=Finalidad sumativa; p2=Finalidad formativa; p3=Comprensión del problema y capacidad comunicativa; p4=Ajuste y evaluación de la solución del problema; p5=Estrategias de evaluación formativa; p6=Exámenes como instrumentos de evaluación; p7=Técnicas e instrumentos de evaluación alternativos; p8=Aspectos cognitivos que componen la calificación; p9=Aspectos no cognitivos que componen la calificación; p10= Comunicación con el colectivo escolar; p11= Comunicación con estudiantes y familias.

Por otra parte, al explorar las diferencias entre los perfiles *intensivo* y *moderado*, en función de las variables sociodemográficas, se tiene que la edad es una variable con diferencias estadísticamente significativas entre los mismos, donde las y los docentes del perfil moderado (M=39.00, DE=9.81) son ligeramente mayores que quienes integran el perfil intensivo (M=37.12; DE=9.52), aunque el tamaño del efecto indica que la diferencia es pequeña (d=0.20, 95% CI [0.14- 0.25]).

En cuanto al sexo, también se encontraron diferencias significativas, puesto que un mayor número de docentes mujeres pertenece al *perfil intensivo*, equivalente a una proporción del 79.67% al interior de este grupo. De igual forma, una mayor proporción de docentes hombres se ubican en el *perfil moderado* (25.43% contra 20.33% en el *perfil intensivo*). No obstante, el tamaño del efecto es básicamente nulo (Tabla 3).

Tabla 3

Frecuencias y resultados de la prueba chi-cuadrado entre perfiles docentes y variables de contexto

Variables	N	Perfil intensivo		Perfil moderado		X ²	gl	p	w																																																																																											
		%	N	%	N																																																																																															
		Sexo	Mujer	2,038	79.67%					1,578	74.57%	17.18	1	<0.001	0.060																																																																																					
	Hombre	520	20.33%	538	25.43%	Institución de formación inicial	Escuela Normal	1,419	55.47%	1,344	63.52%					36.45	2	<0.001	0.090	Universidad pública	517	20.21%	393	18.57%	Universidad privada	622	24.32%	379	17.91%	Último grado de estudios concluido	Licenciatura	2,074	81.8%	1,742	82.33%	5.79	3	0.122	-	Especialidad	76	2.97%	41	1.94%	Maestría	379	14.82%	314	14.84%	Doctorado	29	1.13%	19	0.90%	Participación en espacios de formación continua ^a	Sí	2,135	83.46%	1,583	74.81%	53.29	1	<0.001	0.055	No	423	16.54%	533	25.19%	Años de experiencia	0-5 años	966	37.76%	658	31.10%	48.83	5	<0.001	0.103	6-10 años	436	17.4%	287	13.56%	11-15 años	345	13.49%	343	16.21%	16-20 años	317	12.39%	320	15.12%	21-25 años	302	11.81%	296	13.99%	Más de 26 años
Institución de formación inicial	Escuela Normal	1,419	55.47%	1,344	63.52%		36.45	2	<0.001	0.090																																																																																										
	Universidad pública	517	20.21%	393	18.57%																																																																																															
	Universidad privada	622	24.32%	379	17.91%	Último grado de estudios concluido					Licenciatura	2,074	81.8%	1,742	82.33%	5.79	3	0.122	-	Especialidad	76	2.97%	41	1.94%	Maestría	379	14.82%	314	14.84%		Doctorado	29	1.13%	19	0.90%					Participación en espacios de formación continua ^a	Sí	2,135	83.46%	1,583	74.81%	53.29	1	<0.001	0.055	No	423	16.54%	533	25.19%	Años de experiencia	0-5 años	966	37.76%	658	31.10%	48.83	5	<0.001	0.103	6-10 años	436	17.4%	287	13.56%		11-15 años	345	13.49%	343	16.21%					16-20 años	317	12.39%	320	15.12%	21-25 años	302	11.81%	296	13.99%	Más de 26 años	192	7.51%	212	10.02%						
Último grado de estudios concluido	Licenciatura	2,074	81.8%	1,742	82.33%		5.79	3	0.122	-																																																																																										
	Especialidad	76	2.97%	41	1.94%																																																																																															
	Maestría	379	14.82%	314	14.84%																																																																																															
	Doctorado	29	1.13%	19	0.90%	Participación en espacios de formación continua ^a					Sí	2,135	83.46%	1,583	74.81%	53.29	1	<0.001	0.055	No	423	16.54%	533	25.19%	Años de experiencia	0-5 años	966	37.76%	658	31.10%	48.83	5	<0.001	0.103	6-10 años	436	17.4%	287	13.56%	11-15 años	345	13.49%	343	16.21%	16-20 años	317	12.39%	320	15.12%	21-25 años	302	11.81%	296	13.99%		Más de 26 años	192	7.51%	212	10.02%																																								
Participación en espacios de formación continua ^a	Sí	2,135	83.46%	1,583	74.81%		53.29	1	<0.001	0.055																																																																																										
	No	423	16.54%	533	25.19%	Años de experiencia					0-5 años	966	37.76%	658	31.10%	48.83	5	<0.001	0.103	6-10 años	436	17.4%	287	13.56%		11-15 años	345	13.49%	343	16.21%					16-20 años	317	12.39%	320	15.12%	21-25 años	302	11.81%	296	13.99%	Más de 26 años	192	7.51%	212	10.02%																																																			
Años de experiencia	0-5 años	966	37.76%	658	31.10%		48.83	5	<0.001	0.103																																																																																										
	6-10 años	436	17.4%	287	13.56%																																																																																															
	11-15 años	345	13.49%	343	16.21%																																																																																															
	16-20 años	317	12.39%	320	15.12%																																																																																															
	21-25 años	302	11.81%	296	13.99%																																																																																															
	Más de 26 años	192	7.51%	212	10.02%																																																																																															

Variables	N	Perfil intensivo		Perfil moderado		X ²	gl	p	w
		%	N	%	N				
Grado escolar donde imparte clase	1ro y 2do grado	917	35.85%	693	32.75%	5.98	3	0.113	-
	3ro y 4to grado	758	29.63%	673	31.81%				
	5to y 6to grado	813	31.78%	682	32.23%				
	Multigrado	70	2.74%	68	3.21%				
Sector al que pertenece la escuela en que trabaja actualmente	Público	2,095	81.90%	1,901	89.84%	58.86	1	<0.001	0.118
	Privado	463	18.10%	215	10.16%				
Modalidad de la escuela en la que trabaja según el tipo de estructura organizativa	Organización completa	2,432	95.7%	2,019	95.42%	0.30	1	0.585	-
	Multigrado	126	4.93%	97	4.58%				
Modalidad de la escuela en que trabaja según el tipo de jornada	Jornada regular	2,212	86.47%	1,834	86.67%	0.04	1	0.842	-
	Jornada extendida	346	13.53%	282	13.33%				
Tipo de asentamiento de la escuela	Urbano	1,540	60.20%	1,160	54.82%	13.96	2	<0.001	0.055
	Rural	432	16.89%	396	18.71%				
	Marginado	586	22.91%	560	26.47%				

Nota. N=4,674. Se resalta en negritas las variables que resultaron estadísticamente significativas. Valoración del tamaño del efecto *w* para el resultado de la prueba chi-cuadrado con base en Cárdenas y Arancibia (2014).

^a Se refiere a espacios de formación continua en temas relacionados con la evaluación del aprendizaje en los últimos cinco años.

En la Tabla 3, también se observa que, con respecto a la institución de formación inicial, se presentan diferencias significativas, aunque con tamaños del efecto menores, al ubicarse una mayor proporción de docentes que estudiaron la licenciatura en alguna universidad pública o privada en la composición del *perfil intensivo* que en el *perfil moderado*. Por su parte, los y las docentes que egresaron de escuelas normales¹, aunque

en frecuencias absolutas también se concentran en el *perfil intensivo*, representan una proporción mayor en la integración del *perfil moderado* (63.52%).

Por otra parte, el *perfil intensivo* reunió a un mayor número de docentes que han participado en actividades de formación continua en temas relacionados con la evaluación del aprendizaje en los últimos cinco años. A este respecto, el 83.46% del personal docente que integra el *perfil intensivo* ha participado en cursos, seminarios, talleres, diplomados u otras opciones formativas en esta temática a diferencia de un 74.81% en el *perfil moderado*.

En relación con las variables sociodemográficas de carácter laboral, los años de servicio como docentes de primaria representan diferencias significativas, donde es posible distinguir que en la integración del *perfil intensivo* existe una proporción mayor de docentes de 0 a 10 años que en el *perfil moderado*. De forma contraria, el profesorado de más de 11 años de experiencia conforma en términos proporcionales al *perfil moderado* en mayor medida que al *perfil intensivo*.

Respecto al sector al que pertenece la escuela, si bien un mayor número de docentes tanto de escuelas públicas como privadas conforman el *perfil intensivo*, al interior de cada perfil es posible detectar diferencias. En este sentido, en la conformación del *perfil intensivo* se observa mayor proporción de docentes de escuelas privadas que el *perfil moderado*, a su vez que en el *perfil moderado* se cuenta con una proporción mayor de docentes que trabajan en escuelas públicas.

El tipo de asentamiento de la escuela es una variable que también arrojó diferencias significativas, donde el *perfil intensivo* se compone por una proporción mayor de docentes que trabajan en el contexto urbano y una menor proporción por profesores de zonas rurales o marginadas respecto al *perfil moderado*.

Finalmente, cabe destacar que las variables correspondientes al último grado de estudios concluido, el grado escolar en que se imparte clase, la estructura organizativa de la escuela y el tipo de jornada no presentaron diferencias significativas entre perfiles.

Discusión

Dentro de la muestra de docentes de primaria en Baja California, México, se identificaron dos perfiles asociados a las prácticas de evaluación en matemáticas, denominados como *perfil intensivo* y *perfil moderado*.

El *perfil intensivo* es similar al perfil entusiasta descrito en los estudios de Sayac y Veldhuis (2021), Veldhuis y Van Den Heuvel-Panhuizen (2014) y Zhao et al. (2018), donde este grupo se ubica por encima del promedio en la frecuencia con que realizan las distintas prácticas de evaluación aquí documentadas. Un aspecto para destacar es que, en estos estudios el perfil entusiasta no resultaba el grupo más numeroso en la muestra de docentes, a diferencia de la presente investigación donde el profesorado del perfil intensivo representa poco más de la mitad del profesorado participante.

Por su parte, de acuerdo con los resultados de los estudios referidos, el *perfil moderado* se asocia al perfil poco entusiasta, ya que el personal docente registra puntajes por debajo de la media en las prácticas de evaluación objeto de análisis. Sin embargo,

para fines de este estudio se optó por utilizar el término *moderado*, ya que, si bien los patrones de respuesta registran una menor frecuencia en los comportamientos explorados, estos distan de considerarse poco frecuentes.

Cabe destacar que el grupo de docentes con adscripción al perfil intensivo, no obstante, y refieren una mayor frecuencia en prácticas evaluativas asociadas a un enfoque formativo, también tienden a considerar criterios no cognitivos en el otorgamiento de calificaciones escolares, lo cual se ha documentado como una práctica no recomendable con impactos reducidos en el logro de los aprendizajes esperados. (Chappuis et al., 2012; Dagdag y Dagdag, 2020, Paderes y Ramos, 2024). No obstante, estos resultados refuerzan la importancia de plantear estrategias para apoyar al personal docente a determinar con claridad las calificaciones que asignan a sus estudiantes desde una perspectiva pedagógica que apoye el aprendizaje.

Además, resulta relevante en términos interpretativos que, si bien existen diferencias básicamente en todas las dimensiones involucradas en el proceso evaluativo —finalidad, objeto, estrategias formativas de evaluación, asignación de calificación y comunicación de resultados—, el profesorado no refirió prácticas diferenciadas en cuanto a las herramientas que emplea para recuperar evidencias de aprendizaje de sus estudiantes, al menos en relación con las técnicas e instrumentos incluidos en el cuestionario.

Por otra parte, fue posible caracterizar a los perfiles con base en variables sociodemográficas, si bien hay que tomar estos resultados con cautela, dado que los tamaños del efecto fueron bajos, resulta relevante que siete de estas variables presentaron diferencias significativas. En este sentido, el *perfil intensivo* se compuso proporcionalmente, en comparación con el *perfil moderado*, por más docentes mujeres, de menor edad, egresados(as) de universidades, docentes que han participado en esquemas de formación asociados a la evaluación en los últimos cinco años, con menos de 10 años de experiencia, que trabajan en primarias de sostenimiento privado y localizadas en zonas urbanas.

Al respecto, se observan discrepancias y similitudes con los estudios previos en cuanto a las diferencias encontradas entre perfiles en términos de variables sociodemográficas. Al respecto, en la investigación de Veldhuis y Van Den Heuvel-Panhuizen (2014), desarrollada en Países Bajos, el género, la edad, el número de años de experiencia, la cualificación profesional y el grado escolar en que imparten clase representaron diferencias significativas, lo que coincide con este estudio a excepción del grado escolar. Para el caso de Sayac y Veldhuis (2021), quienes investigaron con docentes de Francia, se encontraron diferencias en algunas variables en función de los años de experiencia, pero de forma contraria la edad, el sexo y los estudios previos no se relacionaron con la asignación a ninguno de los grupos. Finalmente, en Zhao et al. (2018), el género, los años de experiencia también representaron diferencias entre perfiles al igual que en la muestra de profesores de esta investigación.

Los resultados obtenidos conducen a aceptar parcialmente la hipótesis en cuanto a que las variables sociodemográficas permiten caracterizar perfiles docentes asociados a las prácticas de evaluación en matemáticas, lo cual puede orientar la implementación de estrategias diferenciadas para apoyar la labor docente.

Conclusiones

Esta investigación tuvo como objetivo identificar los perfiles docentes asociados a las prácticas de evaluación en matemáticas y su diferenciación en términos de variables sociodemográficas. En una muestra amplia de docentes de primaria en Baja California, México, se encontraron los perfiles intensivo y moderado, con claras diferencias en las distintas dimensiones del proceso evaluativo, con excepción de las técnicas e instrumentos que los docentes refieren aplicar para evaluar a sus estudiantes. A su vez, una serie de variables sociodemográficas como el sexo, la edad, años de experiencia, formación inicial, formación continua, el sostenimiento de la escuela y tipo de asentamiento permitieron diferenciar, aunque con tamaños del efecto bajos, a estos dos grupos de docentes que aplican en mayor o menor medida la evaluación en el aula.

La posibilidad de identificar perfiles de profesorado con prácticas distintas, además de contribuir a la comprensión teórica del objeto de estudio, permite contar con un insumo para el planteamiento de estrategias diferenciadas en apoyo a la práctica docente. Además, en términos metodológicos, el presente estudio demostró que el ACL es una técnica adecuada para abordar objetivos de investigación de esta naturaleza.

A manera de prospectiva, se recomienda considerar en estudios futuros el vincular los perfiles de evaluación con los niveles de desempeño del alumnado, para identificar si una mayor presencia de evaluación en el aula repercute en mayores aprendizajes del alumnado. Asimismo, es deseable recuperar información de otros actores educativos, como estudiantes, familias y autoridades educativas, a fin de trazar un mapa más claro y completo del proceso evaluativo.

Notas.

1. La formación normalista es aquella recibida dentro de una institución formadora de docentes, que en el contexto mexicano se conoce como Escuelas Normales. El personal docente que se ubica en esta categoría ha cursado un plan de estudios enfocado exclusivamente a la enseñanza en el nivel primaria.

Referencias

- Acar-Erdol, T. y Yildizli, H. (2018). Classroom assessment practices of teachers in Turkey. *International Journal of Instruction*, 11(3), 587–602. <https://doi.org/10.12973/iji.2018.11340a>
- Bauer, J. (2022). A Primer to Latent Profile and Latent Class Analysis. En M. Goller, E. Kyndt, S. Paloniemi, y C. Damşa (Eds.), *Methods for Researching Professional Learning and Development: Challenges, Applications and Empirical Illustrations* (pp. 243-268). Springer International Publishing. <https://link.springer.com/book/10.1007/978-3-031-08518-5>
- Black, P. y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31. <https://doi.org/10.1007/s11092-008-9068-5>

- Braund, H. y DeLuca, C. (2018). Elementary students as active agents in their learning: an empirical study of the connections between assessment practices and student metacognition. *Australian Educational Researcher*, 45, 65–85. <https://doi.org/10.1007/s13384-018-0265-z>
- Cárdenas, M. y Arancibia, H. (2014). Potencia estadística y cálculo del tamaño del efecto en G*power: complementos a las pruebas de significación estadística y su aplicación en psicología. *Salud & Sociedad*, 5(2), 210-224. <https://doi.org/10.22199/S07187475.2014.0002.00006>
- Chan, K. T. (2021). Embedding formative assessment in blended learning environment: The case of secondary Chinese language teaching in Singapore. *Education Sciences*, 11(7), 360. <https://doi.org/10.3390/educsci11070360>
- Chappuis, J., Stiggins, R., Chappuis, S., y Arter, J. (2012). *Classroom Assessment for Student Learning. Doing it right—Using it well* (2nd ed.). Pearson.
- Dagdag, J. M. H. y Dagdag, J. D. (2020). Constructivism and the mathematics classroom assessments of elementary teachers. *Journal of Critical Reviews*, 7(12), 816–823. <https://doi.org/10.31838/jcr.07.12.144>
- Demosthenous, E., Christou, C., y Pitta-Pantazi, D. (2021). Mathematics classroom assessment: A framework for designing assessment tasks and interpreting students' responses. *European Journal of Investigation in Health, Psychology and Education*, 11(3), 1088–1106. <https://doi.org/10.3390/ejihpe11030081>
- Ebel, R. L. y Frisbie, D.A. (1986). *Essentials of Education Measurement* (4th ed.). Prentice-Hall.
- George, D. y Mallery, P. (2003). *SPSS for Windows step by step: A Simple Guide and Reference* (4th ed.). Allyn y Bacon.
- Herbert, S. (2021). Overcoming Challenges in Assessing Mathematical Reasoning. *Australian Journal of Teacher Education*, 46(8), 17–30. <https://doi.org/10.14221/ajte.2021v46n8.2>
- INEE (2018). *Planea Resultados nacionales 2018, 6to primaria. Lenguaje y comunicación, Matemáticas*. Plan Nacional de Evaluación de los aprendizajes, Instituto Nacional para la evaluación de la Educación. https://www.inee.edu.mx/images/stories/2018/planea/PLANEA06_Rueda_de_prensa_27nov2018.pdf
- Lara, L. y Saracostti, M. (2019). Effect of parental involvement on children's academic achievement in Chile. *Frontiers in Psychology*, 10, 1464. <https://doi.org/10.3389/fpsyg.2019.01464>
- Leenknecht, M., Wijnia, L., Köhler, M., Fryer, L., Rikers, R., y Loyens, S. (2021). Formative assessment as practice: the role of students' motivation. *Assessment and Evaluation in Higher Education*, 46(2), 236–255. <https://doi.org/10.1080/02602938.2020.1765228>
- Linzer, D. A. y Lewis, J. B. (2011). poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, 42(10), 1-29. <https://www.jstatsoft.org/v42/i10/>
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151–1169. <https://doi.org/10.6018/analesps.30.3.199361>

- Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. En T. D. Little (Ed.), *The Oxford handbook of quantitative methods, Vol. 2: Statistical Analysis* (pp. 551–611). Oxford University Press.
- Montoya, O. (2007). Aplicación del análisis factorial a la investigación de mercados. Caso de estudio. *Scientia Et Technica*, 13(35), 281-286. <https://www.redalyc.org/pdf/849/84903549.pdf>
- Navarro-Mosquera, N. G., Falconí-Asanza, A. V., y Espinoza-Cordero J. (2017). El mejoramiento del proceso de evaluación de los estudiantes de la educación básica. *Universidad y Sociedad*, 9(4), 58-69. http://scielo.sld.cu/scielo.php?script=sci_abstractypid=S2218-36202017000400008yIngl=esynrm=iso
- Nylund-Gibson, K. y Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, 4(4), 440- 461. <https://doi.org/10.1037/tps0000176>
- Ondé, D. y Alvarado, J. M. (2019). Análisis de clases latentes como técnica de identificación de tipologías. *Revista INFAD de Psicología. International Journal of Developmental and Educational Psychology*, 5(1), 251–260.. <https://doi.org/10.17060/ijodaep.2019.n1.v5.1641>
- Organización para la Cooperación y el Desarrollo Económico (OCDE) (2023). *PISA 2022 Results (Vol I and II) - Country Notes: México*. <https://www.oecd.org/publication/pisa-2022-results/country-notes/mexico-519eaf88/#chapter-d1e11>
- Paderes, B. M. y Ramos, C. (2024). Teachers' Decision-Making in Grading Students at Risk of Academic Failure during the Implementation of the Basic Education Learning Continuity Plan. *FEU Journal of Graduate Students' Research*, 1(1), 1-17. <https://www.feu.edu.ph/wp-content/uploads/2024/02/Grade-Decision-Making-Paderes-Ramos-2023.pdf>
- Polya, G. (1965). *Cómo plantear y resolver problemas*. (J. Zugazagoitia, Trad.). Editorial Trillas.
- Ravela, P., Picaroni, B., y Loureiro, G. (2017). *¿Cómo mejorar la evaluación en el aula? Reflexiones y propuestas de trabajo para docentes* (2a ed.). Grupo Magro Editores.
- Sayac, N. y Veldhuis, M. (2021). Mathematics Assessment Practices of Primary School Teachers in France. *International Journal of Science and Mathematics Education*, 20, 1595-1610. <https://doi.org/10.1007/s10763-021-10229-3>
- Schellekens, L. H., Bok, H., de Jong, L. H., van der Schaaf, M. F., Kremer, W., y Van der Vleuten, C. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, 71, 101094. <https://doi.org/10.1016/j.stueduc.2021.101094>
- Sheikh, A. B. M. y Manap, M. N. A. (2024). Classroom Assessment Practices. *Open Journal of Social Sciences*, 12(3), 239-258. <https://doi.org/10.4236/jss.2024.123018>
- Schildkamp, K., Van der Kleij, F. M., Heitink, M. C., Kippers, W. B., y Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for

- classroom practice. *International Journal of Educational Research*, 103, 101602. <https://doi.org/10.1016/j.ijer.2020.101602>
- Tariq, E., Asif, A., y Nazirullah. (2023). Assessment Practices and Assessment Skills of Teachers at Secondary School Level: Factorability and Associative Perspective. *Journal of Peace, Development and Communication*, 7(01), 101–116. <https://doi.org/10.36968/JPDC-V07-I01-10>
- Veldhuis, M. y Van Den Heuvel-Panhuizen, M. (2014). Primary school teachers' assessment profiles in mathematics education. *PLoS ONE*, 9(1), e86817. <https://doi.org/10.1371/journal.pone.0086817>
- Wafubwa, R. N. y Csikos, C. (2021). Formative Assessment as a Predictor of Mathematics Teachers' Levels of Metacognitive Regulation. *International Journal of Instruction*, 14(1), 983–998. <https://doi.org/10.29333/IJI.2021.14158A>
- Weller, B. E., Bowen, N. K., y Faubert, S. J. (2020). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 46(4), 287–311. <https://doi.org/10.1177/0095798420930932>
- Zhao, X., Van den Heuvel-Panhuizen, M., y Veldhuis, M. (2018). Chinese Primary School Mathematics Teachers' Assessment Profiles: Findings from a Large-Scale Questionnaire Survey. *International Journal of Science and Mathematics Education*, 16(7), 1387–1407. <https://doi.org/10.1007/s10763-017-9841-3>
- Zulliger, S., Buholzer, A., y Ruelmann, M. (2022). Observed Quality of Formative Peer and Self-Assessment in Everyday Mathematics Teaching and its Effects on Student Performance. *European Journal of Educational Research*, 11(2), 663–680. <https://doi.org/10.12973/eu-jer.11.2.663>

Fecha de recepción: 4 febrero, 2024.

Fecha de revisión: 13 marzo, 2024.

Fecha de aceptación: 12 de noviembre, 2024