



ISSN: 1989-0397

**Revista  
Iberoamericana  
de Evaluación Educativa**

Mayo 2018 - Volumen 11, número 1  
doi: 10.15366/rie2018.11.1

[rinace.net/rie/  
revistas.uam/rie](http://rinace.net/rie/revistas.uam/rie)



Red Iberoamericana  
de Investigación  
sobre Cambio  
y Eficacia Escolar

# CONSEJO EDITORIAL

## DIRECTOR

F. Javier Murillo

## EDITORA

Nina Hidalgo Farran

## ASESOR EDITORIAL

Manuel Lorite Becerra

## CONSEJO DIRECTIVO

Marcela Gajardo (Programa de Promoción de la Reforma Educativas de América Latina y El Caribe, PREAL)

Sergio Martinic (Pontificia Universidad Católica de Chile)

Carlos Pardo (Instituto Colombiano para la Evaluación de la Educación, ICFES)

Margarita Poggi (Instituto Internacional de Planeamiento de la Educación -IIFE-. UNESCO, Argentina)

Francisco Soares (Universidade Federal de Minas Gerais, Brasil)

## CONSEJO CIENTÍFICO

Juan Manuel Álvarez. Universidad Complutense de Madrid, España.

Patricia Arregui. Grupo de Análisis para el Desarrollo (GRADE), Perú.

Daniel Bogoya. Universidad Pedagógica Nacional, Colombia.

Nigel Brooke. Universidade Federal de Minas Gerais, Brasil.

Leonor Cariola. Ministerio de Educación, Chile.

María do Carmo Clímaco. Universidade Lusófona de Humanidades e Tecnologias (ULHT), Portugal.

Cristian Cox. Pontificia Universidad Católica de Chile.

Santiago Cueto. Grupo de Análisis para el Desarrollo (GRADE).

Tabaré Fernández. Universidad de la República, Uruguay.

Juan Enrique Froemel. Universidad UNIACC, Chile.

Reyes Hernández-Castilla. Universidad Autónoma de Madrid, España.

Rubén Klein. Fundação Cesgranrio, Brasil.

Luis Lizasoain. Universidad del País Vasco/Euskal Herriko Unibertsitatea, España.

Jorge Manzi. MIDE-UC, Pontificia Universidad Católica de Chile.

Joan Mateo. Universidad de Barcelona, España.

Liliana Miranda. Ministerio de Educación de Perú.

Carlos Muñoz Izquierdo. Universidad Iberoamericana Ciudad de México, México.

Margarita Peña. Instituto Colombiano para la Evaluación de la Educación, ICFES.

Dagmar Raczynski. Asesorías para el Desarrollo, Chile.

Héctor Rizo. Universidad Autónoma de Occidente, Colombia.

Mario Rueda. Universidad Nacional Autónoma de México.

Guadalupe Ruíz. Universidad Autónoma de Aguascalientes, México.

Ernesto Schiefelbein. Universidad Autónoma de Santiago, Chile.

Alejandra Schullmeyer. Instituto Nacional de Estudios Pedagógicos, Brasil.

Javier Tejedor. Universidad de Salamanca, España.

Flavia Terigi. Universidad de Buenos Aires, Argentina.

Alexander Ventura. Universidade de Aveiro, Portugal.

# ÍNDICE

## Editorial

- ¿Ya han pasado 10 años? ¡Pero si nos queda todo por hacer!** 4  
*F. Javier Murillo y Nina Hidalgo*

## Temática Libre

- Evaluación Abreviada del Conocimiento Fonológico mediante la Prueba P.E.C.O.** 9  
*José Luís Ramos Sánchez y Ana Isabel González Contreras*
- Construcción y Validación de Instrumentos para la Evaluación de la Práctica Pedagógica en Educación Básica del Estado de Yucatán** 23  
*Omar Eduardo Kú Hernández y Wilson Jesús Pool Cibrián*
- Análisis sobre la Evaluación de la Calidad Educativa en América Latina: Caso Colombia** 43  
*Deivis Robinson Mosquera Albornoz*
- Estilos de Enseñanza y Representaciones sobre Evaluación y Aprendizaje** 57  
*Claudia Alejandra Mazzitelli, Ana María Guirado y María Julieta Laudadio*
- Efecto de la Integración de las Artes Visuales en el Desarrollo de la Lectura y la Escritura** 73  
*Wanda I. Padilla Martínez y Jaime Ortiz-Vega*
- Influencia de la Longitud, la Complejidad y la Inferencia en la Resolución de Pruebas de Comprensión Lectora Inferencial** 95  
*Alejandra Platas-García, J. Martín Castro-Manzano, Verónica Reyes-Meza e Irene Gaona-Gordillo*
- Aplicación del Método Delphi Modificado para la Validación de un Cuestionario de Incorporación de las TIC en la Práctica Docente** 113  
*Carlos Enrique George Reyes y Laura Trujillo Liñan*
- La Experiencia de la Evaluación Docente en México: Análisis Crítico de la Imposición del Servicio Profesional Docente** 135  
*Francisco Guzmán Marín*
- Una Propuesta de Ejes para el Análisis del Uso de Resultados de Evaluaciones a Gran Escala del Aprendizaje por Docentes y Directores** 159  
*Guadalupe Ruiz Cuéllar y Paloma Prado Robledo*

## Editorial:

# ¿Ya han Pasado Diez Años? ¡Pero si nos Queda Todo por Hacer!

## Has It Been Ten Years Already?? But we have everything left to do!

F. Javier Murillo \*  
Nina Hidalgo

Universidad Autónoma de Madrid - España

Perece mentira, ya han pasado 10 años. En mayo de 2008 se publicó el primer número de la Revista Iberoamericana de Evaluación Educativa y con este ya vamos por el 25. El título del primer editorial ya era toda una declaración de intenciones: “Evaluación educativa como derecho Humano”. El argumento del mismo es sencillo solo con la evaluación es posible garantizar el derecho humano a revivir una educación de calidad; pero para ello la evaluación debe también ser de calidad. En esa línea:

*La Revista Iberoamericana de Evaluación Educativa nace con ese propósito: contribuir a mejorar la calidad de la Evaluación Educativa que se realiza en Iberoamérica de tal forma que logre incrementar significativamente los niveles de calidad y equidad educativa. Y ello mediante, el fomento de la investigación, el análisis y la reflexión; la generación de redes de estudiosos sobre el tema; y la difusión de estudios, aportaciones teóricas y buenas prácticas sobre el tema. (Murillo y Román, 2008, p 3).*

Diez años después de escribir estas líneas no nos parece descabellado afirmar que estamos consiguiendo el objetivo propuesto: hemos hecho una humilde pero real contribución a la mejora de la Evaluación Educativa en Iberoamérica; ojalá que ello se haya traducido en una mayor calidad y equidad educativa. De lo que estamos seguros es que, si algo hemos logrado, todo nos queda por conseguir.

En este tiempo se han publicado 310 artículos, lo que significa 6.000 páginas, que han sido citados, a fecha de hoy, por casi 3.000 textos<sup>1</sup>. Como luego veremos, la Revista está indexada en 27 bases de datos, directorios o catálogos. Ni mucho ni poco, es lo que es.

De los que era la revista en sus inicios a lo que ahora es, queda mucho y mucho ha variado. Seguramente son tres los acontecimientos que más definieron esa evolución. En primer lugar, la marcha de Marcela Román como codirectora de la Revista hace ya casi 4 años. Sin ella la revista no solo sería mucho peor, seguro que ni hubiera existido. Todo nuestro reconocimiento y agradecimiento. También fue relevante el cambio de Editora

---

<sup>1</sup> <https://scholar.google.com/citations?user=LghEe7oAAAAJ&hl=en>

en el año 2013, Nina Hidalgo asumió las funciones que tan fantásticamente ejerció Verónica González de Alba desde la creación de la Revista. El tercero hito, mucho más prosaico, es que en 2015 la Revista se empezó a publicar también en el sistema OJS de la Universidad Autónoma de Madrid, cosa que nos permitió, por ejemplo, contar con el ya imprescindible DOI<sup>2</sup>. Así, ahora la revista aparece tanto en el formato tradicional en la página de RINACE<sup>3</sup> como en OJS en el portar de revistas de la UAM<sup>4</sup>.

En esta década, se han publicado 310 artículos organizados en 24 números, de los cuales, 20 han sido los números semestrales que la revista presenta puntualmente en mayo y noviembre de cada año (226 artículos) y cuatro números especiales dedicados a recoger las memorias de diferentes Coloquios celebrados de la Red Iberoamericana de Investigadores sobre Evaluación de la Docencia (84 artículos).

Analizando los 20 números regulares publicados, 15 de ellos han tenido una sección temática que combinaba la publicación de artículos temáticos con otros cinco que eran de tema “libre”. De esta forma, de los 226 trabajos publicados, casi un 70% ha conformado han correspondido a secciones temáticas. En ellas, se han tratado una miríada de temas y enfoques de la evaluación educativa: de resultados, de escuelas, de programas y políticas, de docentes universitarios y no universitarios, del liderazgo escolar, de educación inicial y preescolar, de la educación rural, de TIC en educación, de experiencias en didácticas específicas, de materiales educativos de prácticas de intercambio y convivencia escolar, de evaluación para la Justicia Social... tantos temas abordados y tantos por abordar. Gracias a todos y cada uno de los coordinadores y coordinadoras de los números, su trabajo ha sido fundamental para llegar hasta aquí.

Los cuatro números especiales han sido peticiones especiales para la publicación de 84 trabajos de cuatro ediciones del Coloquio Iberoamericano de la Red Iberoamericana de Investigadores sobre Evaluación de la Docencia (RIIED). En estos cuatro coloquios se han abordado temas centrados en la Evaluación de la Docencia tanto Universitaria como no Universitaria, así como los avances y retos que le queda a la evaluación de la práctica del profesorado.

El 62% de los trabajos que conforman los diferentes números de la Revista Iberoamericana de Evaluación Educativa son investigaciones empíricas, el 22% artículos de ensayo o reflexión teórica, así como estados del arte sobre temas concretos de evaluación, y el 19% de los artículos restantes están dedicados a experiencias o innovaciones sobre evaluación educativa. Si analizamos la evolución de los artículos publicados, podemos ver que cada vez se publican más artículos empíricos en contraposición a las innovaciones o experiencias, las cuales han ido decreciendo hasta no publicar ninguna en los últimos cuatro años (gráfico 1).

---

<sup>2</sup> <https://doi.org/10.15366/riee>

<sup>3</sup> <http://www.rinace.net/riee/>

<sup>4</sup> <https://revistas.uam.es/riee>

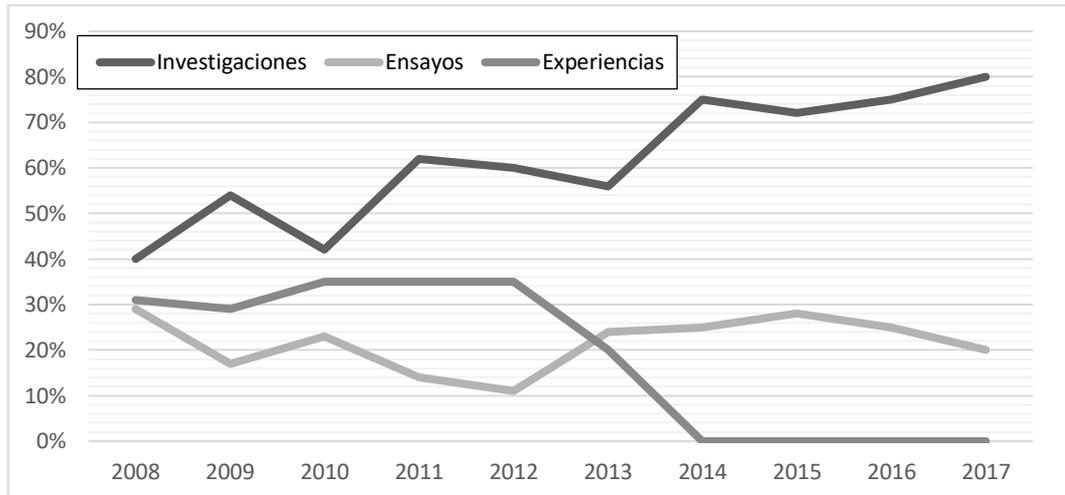


Gráfico 1. Evolución de la distribución de los artículos publicados en la Revista Iberoamericana de Evaluación Educativa por tipo

Fuente: Elaboración propia.

Uno de los logros de la Revista es haber recibido y publicado trabajos de la gran mayoría de países de América Latina, aunque su distribución es muy desigual entre ellos (grafico 2). Efectivamente, casi el 28% proceden de México, seguidos por los de España, con el 20% de los trabajos publicados, y Chile con el 16%. A partir de ahí una larga lista de países: Brasil, Colombia, Argentina, Perú, Portugal, Cuba, Paraguay, Venezuela, Honduras, Costa Rica y Ecuador. También hay algunos artículos publicados por autores y autoras procedentes de fuera de la Región: de Estados Unidos, de Canadá, de Inglaterra o de Nueva Zelanda. El reto es recibir buenos artículos de los países ausentes de Centroamérica y el Caribe: Nicaragua, El Salvador, República Dominicana o Guatemala.

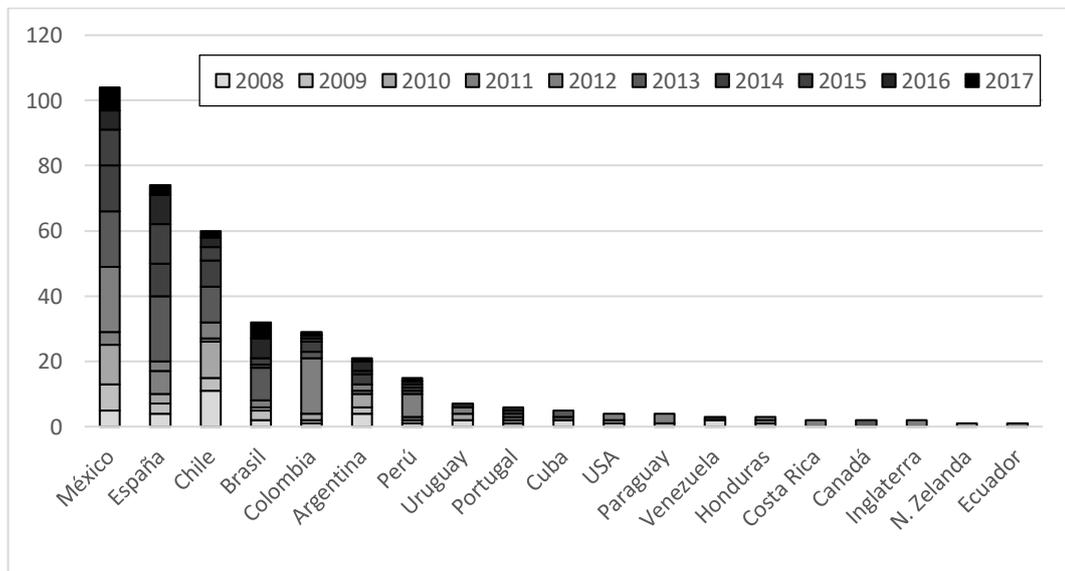


Gráfico 2. Nacionalidad de los autores que han publicado en la Revista Iberoamericana de Evaluación Educativa

Fuente: Elaboración propia.

Estos 10 años han supuesto también un avance en la indexación de la revista en distintas bases de datos. En estos años, la revista ha sido indexada en un total de 21 bases de datos, directorios de evaluación, sistemas de evaluación, servicios de información y catálogos colectivos (cuadro 2). Esto supone una mayor visibilidad y difusión de la revista, y con ello, un mayor número de lectores que se acercan a la publicación.

Cuadro 2. Indexaciones de la Revista Iberoamericana de Evaluación Educativa

TIPO DE INDEXACIÓN	RECURSO
<b>BASES DE DATOS ESPECIALIZADAS</b>	EbscoHost
<b>BASES DE DATOS MULTIDISCIPLINARES</b>	Google Académico Difusión de Alertas en la Red ISOC- Base de datos del Consejo Superior de Investigaciones Científicas de España
<b>DIRECTORIOS</b>	DOAJ- Directory of Open Access Journals Dulcinea Derechos de copyright y las condiciones de auto-archivo de revistas científicas españolas The Elektronische Zeitschriftenbibliothek- Electronic Journals Library E-revistas
<b>SISTEMAS DE EVALUACIÓN</b>	CIRC- Clasificación integrada de revistas científicas DICE - Difusión y Calidad Editorial de las Revistas Españolas de Humanidades y Ciencias Sociales y Jurídicas Latindex/Catálogo -Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal
<b>MODELO DE IDENTIFICACIÓN Y EVALUACIÓN DE REVISTAS</b>	RESH- Revistas Españolas de Ciencias Sociales y Humanidades
<b>SERVICIOS DE INFORMACIÓN</b>	Biblioteca Digital OEI Índice de Revistas de Educación Superior e Investigación Educativa REDIAL- Red Europea de Información y Documentación sobre América Latina REDIB- Red Iberoamericana de Innovación y Conocimiento Científico
<b>CATÁLOGOS COLECTIVOS</b>	Biblioteca Nacional de España WorldCat REBUIN- Red de Bibliotecas Universitarias Catalogue du Système Universitaire de Documentation

Fuente: Elaboración propia.

Todo esto ha de tener reflejo en el impacto de la Revista. La información que nos arroja el Google Scholar es reveladora. A día de hoy (mañana seguro que será más), la Revista tiene un total de 2.964 citas, un índice h de 28 y un índice i10 de 82. Como podemos observar en gráfico 3, desde su nacimiento se han aumentado de una forma muy clara las citas que recibe, aunque también hay que decir que el año 2017 está suponiendo un parón respecto a la espectacular subida. Esperamos que sea solo un tropiezo y continuar una clara tendencia al alza (gráfico 3).

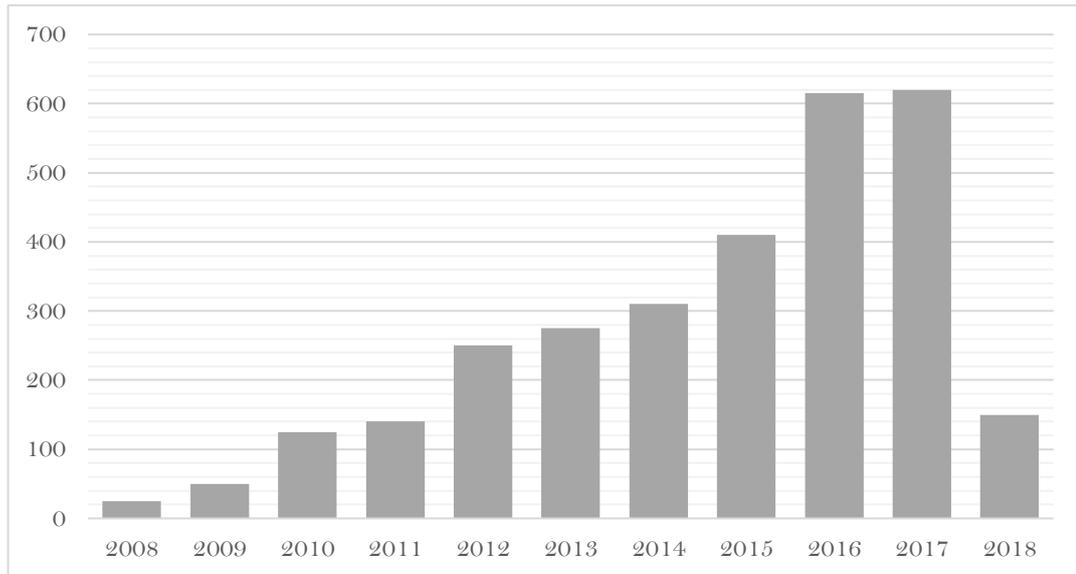


Gráfico 3. Número de citas en la Revista Iberoamericana de Evaluación Educativa  
Fuente: Google Scholar (<https://scholar.google.es/citations?user=LghEe7oAAAAJ&hl=es>)

A pesar de todos los avances realizados, aún queda muchos retos por cumplir. Así, el primero, y más importante, es seguir mejorando la revista en cuanto a su calidad, impacto y bases de datos donde esté indexada. En la actualidad estamos trabajando para que sea reconocida en otras bases de datos más ambiciosas y exigentes. Ojalá que pronto podamos celebrar que lo vamos consiguiendo.

Pero esa no es la meta. La meta la formulamos al inicio de esta reflexión: “contribuir a mejorar la calidad de la Evaluación Educativa que se realiza en Iberoamérica de tal forma que logre incrementar significativamente los niveles de calidad y equidad educativa”. Y aún estamos muy muy lejos de lograrlo.

Conseguir una educación de calidad con equidad que contribuya a una sociedad más justa e inclusiva es nuestra razón de ser. Y no pararemos hasta lograrlo.

## Referencias

Murillo, F. J. y Román, M. (2008). La evaluación educativa como derecho humano. *Revista Iberoamericana de Evaluación Educativa*, 1(1), 1-5.

## Evaluación Abreviada del Conocimiento Fonológico mediante la Prueba P.E.C.O.

### Short Assessment Test by Phonological Awareness P.E.C.O.

José Luís Ramos Sánchez \*  
Ana Isabel González Contreras

Universidad de Extremadura - España

El objetivo de este trabajo es obtener una versión abreviada del conocimiento fonológico a partir de la prueba P.E.C.O. manteniendo una validez adecuada. Entendemos por conocimiento fonológico la habilidad para tomar conciencia y manipular los elementos más simples del lenguaje oral como son las sílabas (conocimiento silábico) y los fonemas (conocimiento fonémico). Un adecuado desarrollo del conocimiento fonológico junto con otras habilidades como la velocidad de denominación son fundamentales para la adquisición del principio alfabético, y por tanto para aprender a leer y a escribir en un idioma transparente como el español. Para llevar a cabo el estudio, se toma la muestra original de tipificación (N=240) formada por alumnos y alumnas de habla española de educación infantil de 5 años (mediana de edad de 5 años y 9 meses) y se utiliza el análisis de regresión múltiple (método de pasos sucesivos) para obtener un modelo de regresión que maximiza una predicción ajustada con nueve ítems. A partir de los resultados se obtuvo una versión abreviada y representativa de la prueba original, que contiene ítems variados en cuanto al tipo de tareas (identificación, adición y omisión) con garantías de predecir el rendimiento en conocimiento fonológico de los alumnos en tres niveles: alto, medio y bajo.

**Palabras clave:** Aprendizaje inicial de la lectoescritura; Análisis de regresión múltiple; Evaluación del conocimiento fonológico; Principio alfabético.

The objective of this work is to obtain an abridged version of phonological awareness from the test P.E.C.O. maintaining adequate validity. We understand phonemic awareness skills to become aware and manipulate the simplest elements of spoken language such as syllables (syllabic knowledge) and phonemes (phonemic awareness). Proper development of phonological awareness along with other skills such as naming speed are essential for the acquisition of the alphabetic principle, and thus to learn to read and write in a transparent language such as Spanish. To conduct the study, the original sample typing (N = 240) formed by students from Spanish-speaking child education of 5 years (median age of 5 years and 9 months) is taken and used analysis multiple regression (stepwise method) to obtain a regression model that maximizes an adjusted prediction nine items. From the results an abbreviated and representative version of the original test, containing various items in the type of tasks (identification, addition and omission) guarantees to predict performance in phonological awareness of students in three levels was obtained: high, medium and low.

**Keywords:** Initial literacy learning; Multiple regression analysis; Evaluation of phonological awareness; Alphabetic principle.

## 1. Introducción

El conocimiento fonológico o conciencia fonológica es la habilidad para tomar conciencia y manipular los elementos más simples del lenguaje oral como son las sílabas (conocimiento silábico) y los fonemas (conocimiento fonémico), y son numerosos los trabajos que demuestran la importancia de esta habilidad metalingüística, tanto para la adquisición de la lectoescritura en educación infantil como variable deficitaria y susceptible de refuerzo en el alumnado disléxico y con dificultades en el aprendizaje lectoescritor. Un buen ejemplo de ello son los trabajos ya clásicos de Sebastián y Maldonado (1987), Sánchez, Rueda y Orrantía (1989), Carrillo y Sánchez (1991, 1996), Domínguez (1992, 1996), Ortíz (1994), Jiménez y Ortiz (1995), Wagner, Torgesen y Rashotte (1994), Ramos (2000, 2004), Ramos y Cuadrado (2004), etc.

Se piensa que el conocimiento fonológico es crucial para llevar a cabo las Reglas de Conversión Grafema-Fonema (RCGF) que son necesarias durante la decodificación fonológica (Manis et al., 1999; Vellutino et al., 2004). Y se ha hecho mucho hincapié en demostrar cómo el trabajar con fonemas es una herramienta específica para la adquisición del principio alfabético (Byrne y Fielding-Barnsley, 1989; De Jong y Olson, 2004). Aunque también sabemos, que una vez que el alumno se ha iniciado en los procesos de decodificación, la influencia del conocimiento fonológico es menor, y la prueba de ello es que los coeficientes de correlación entre el conocimiento fonológico y el rendimiento lectoescritor disminuyen al año siguiente de iniciarse al alumnado en la lectoescritura (ver cuadros 8 y 9 de este trabajo).

Por estas evidencias, el conocimiento fonológico forma parte tanto del conjunto de variables en el proceso de evaluación como de la intervención en gran parte del alumnado con dificultades en los procesos fonológicos de lectura y escritura. La *Prueba de Evaluación del Conocimiento Fonológica* (PECO) constituye un ejemplo contrastado de evaluación del conocimiento fonológico y ha sido utilizada para la evaluación del alumnado y en investigaciones sobre el tema. Pero a veces las características de los sujetos (edad, nivel de atención, etc.) aconsejan, por un motivo u otro, brevedad en la aplicación.

PECO original es una prueba de aplicación individual compuesta por 30 ítems, y por la experiencia en la aplicación de la misma, el tiempo empleado es de 30 a 40 minutos aproximadamente. A veces un tiempo excesivo para algunos alumnos, especialmente para aquellos más predispuestos al cansancio y al abandono de la tarea: alumnado menor de seis años, alumnado con déficit de atención, etc. Y en otras ocasiones, la limitación de tiempo o de presupuesto dificulta la tarea de investigación.

Es cierto que, además del conocimiento fonológico, existen evidencias sobre otras variables que influyen en el aprendizaje inicial de lectura y escritura; por ejemplo, la velocidad de denominación o de nombrado, entendida como el tiempo empleado por el alumno en identificar un estímulo visual. En las tareas típicas de denominación rápida se pide nombrar en voz alta diversos estímulos (dibujos, letras, números o colores), tan rápido como sea posible y se toma como parámetro el tiempo total que se emplea para nombrarlos. Existe correlación entre la velocidad de denominación y el aprendizaje inicial de la lectoescritura en los niños con desarrollo normal (Cutting y Denckla, 2001; de Jong y Van der Leij, 1999; Kirby et al., 2003; Powell et al., 2007).

El estudio realizado por Guzmán et al. (2004) demostró que el déficit de velocidad de denominación en niños con dificultades lectoras es causado por su déficit en el procesamiento fonológico, tomando la idea de que si las tareas de velocidad de denominación realizan un papel importante en la adquisición de representaciones ortográficas de las palabras, sus efectos se deben también considerar en medidas de exactitud y de velocidad. Los resultados demostraron que los alumnos con dificultades de adquisición en la lectura y déficit fonológico eran más lentos en las tareas de nombrar colores, dibujos, letras y números que los lectores normales de su misma edad.

Igualmente, López-Escribano (2007) confirma otros estudios en los que se afirma que la velocidad de denominación contribuye fuertemente a la velocidad de la lectura y al reconocimiento ortográfico. La velocidad de denominación afecta sobre todo a la realización de las tareas de la lectura que requieren una respuesta fluida. Su investigación confirma los resultados que asocian el reconocimiento ortográfico a las habilidades de la lectura.

Por su parte, Gómez et al. (2010) comprobaron que la denominación de letras predijo mejor la ejecución lectora e identificó correctamente a 63% de los niños de 7 años en su estudio (N=121) que posteriormente presentaron dificultades en la velocidad para leer, la cual se ha considerado como el rasgo distintivo de la dislexia en español. Algunas tareas de conciencia fonológica contribuyeron a explicar la eficiencia y la comprensión lectora. Sin embargo, los niños con un *doble déficit* (menor velocidad de denominación y de conciencia fonológica) presentaron el peor rendimiento lector.

Esta conclusión tiene una gran importancia para la detección temprana de las dificultades de la lectoescritura, en el sentido de que sería necesario incluir la evaluación de la velocidad de denominación junto con el conocimiento fonológico al finalizar la educación infantil como predictores de las dificultades de lectura al inicio de la educación primaria o al finalizar la educación infantil.

Nuestro trabajo tiene como objetivo identificar los ítems más relevantes de PECO para predecir lo mejor posible el nivel de conocimiento fonológico del alumno en base a esta prueba. Por tanto, se pretende obtener una versión abreviada de PECO teniendo en cuenta la composición heterogénea del conocimiento fonológico, con objeto de disminuir el tiempo de evaluación manteniendo una adecuada calidad técnica.

## 2. Fundamentación teórica

El conocimiento fonológico es objeto de investigación desde hace más de tres décadas, y son numerosas las investigaciones que utilizan esta variable, tanto en el proceso de explicación causal y relacional de las dificultades lectoescritoras como en el proceso de intervención para mejorar su desarrollo o con la finalidad de prevenir dificultades. A modo de ejemplo, los trabajos de Trias, Cuadro y Costa (2009), Defior y Serrano (2011), Muñoz (2011), Bravo, Villalón y Orellana (2011), Leal y Suro (2012), Suárez, García y Cuetos (2013), Delgado y Sancho (2014), Feld (2014) y González, Cuetos, Vilar y Uceira (2015), son una pequeña muestra del interés actual sobre la importancia de disponer de instrumentos que evalúen el conocimiento fonológico.

Precisamente, dada la necesidad de que los profesionales de la educación y la psicología dispusieran de una prueba para evaluar el desarrollo del conocimiento fonológico, fue

elaborada PECO de Ramos y Cuadrado (2006a), y consecuentemente también fueron elaborados materiales para la intervención (Ramos y Cuadrado, 2006b).

En la elaboración de PECO se tuvo en cuenta que el conocimiento fonológico no tiene una entidad homogénea, sino que depende de distintas variables relacionadas con la estructura de las palabras. Así, PECO fue elaborada ajustándose a los cuatro criterios que más se repiten en la evaluación del conocimiento fonológico. A saber, la unidad fonológica (sílabas y fonemas), el tipo de tarea (análisis y síntesis), la posición del segmento fonológico (inicial, medio y final) y la estructura silábica utilizada.

En cuanto a la *unidad fonológica*, identificamos tres niveles de conocimiento fonológico: palabras, sílabas y fonemas. Las unidades más importantes en la adquisición inicial de la lectoescritura son la fonémica y la silábica por este orden. Son más fáciles las tareas relacionadas con las palabras y las sílabas que con los fonemas.

Según el *tipo de tarea*, se identifican básicamente dos tipos: análisis y síntesis. Las tareas de análisis tienen como objetivo tomar conciencia y manipular unidades aisladas de las palabras tales como las de rima, identificación, aislamiento, recuento, omisión, sustitución e inversión. Las tareas de síntesis conllevan la habilidad para recomponer, recodificar o formar una palabra que se presenta dividida en partes. Este es el objetivo de las tareas de adición y unión de sílabas y/o fonemas. En cuanto a la dificultad, las tareas de rimas e identificación son más fáciles que las de omisión y adición; mientras que las tareas más complejas son las de inversión, que requieren un gran esfuerzo cognitivo no tan necesario para aprender a leer y escribir (Yopp, 1988).

La *posición del segmento fonológico* constituye otra fuente de variabilidad de las tareas, según sea al inicio, en el medio o al final de la palabra. Los resultados del trabajo de Domínguez y Clemente (1993) demuestran que la enseñanza de las habilidades de análisis de las palabras es igual de efectiva utilizando sonidos iniciales, medios o finales, puesto que el aprendizaje por parte de los alumnos en una de las posiciones se generaliza a otras posiciones de la palabra. Aunque admiten que es más fácil reconocer sílabas y fonemas en posición inicial que al final y ambas actividades son más fáciles que identificar sílabas o fonemas en el medio de las palabras. Mientras que, en las tareas de omisión y adición, es más sencillo operar con segmentos finales que iniciales, y ambas son más fáciles que las tareas con unidades mediales.

Finalmente, otra variable clasificatoria de las tareas de conocimiento fonológico es la referida al *tipo de sílaba*. No es lo mismo operar con grupos consonánticos complejos del tipo CCV (Consonante+Consonante+Vocal) o CCVC (ejemplo: **b**razo, **t**ronco, ...) que utilizar sílabas simples, siendo estas últimas más fáciles de trabajar (CV -**g**ato- o VC -**a**lto-) según demuestra la investigación de Treiman (1991).

### 3. Método

Como muestra utilizamos la original de la prueba compuesta por 240 alumnos (n=120) y alumnas (n=120) sin necesidades educativas especiales de educación infantil, de habla española, con un intervalo de edad entre 5 años y 4 meses y 6 años y 4 meses (mediana de 5 años y 9 meses), procedentes de centros educativos de zona urbana y rural, públicos y concertados.

Para llevar a cabo el objetivo de investigación, se utiliza una metodología correlacional a través de la técnica del análisis de regresión múltiple del paquete estadístico SPSS (versión 18). Este tipo de análisis tiene dos finalidades en nuestro estudio. Por un lado, predecir la puntuación total en PECO; y por otro, identificar qué variables son las más relevantes en la predicción. Precisamente estas variables serán las que formen parte de la versión abreviada si se cumplen las condiciones de aplicación y la validación del modelo de regresión.

En cuanto a las variables, las predictoras (independientes) son los 30 ítems que conforman la prueba original y como variable criterio (dependiente) se utilizó la puntuación total en la prueba. Y en relación con el método para la introducción de las distintas variables en el análisis de regresión, se utiliza el de *pasos sucesivos*.

#### 4. Resultados

Presentamos un resumen de la aplicación de la técnica del análisis de regresión lineal multivariada de SPSS (v.18), aunque por brevedad no se incluyen los resultados de las distintas etapas. En el cuadro 1 se muestran las nueve variables incluidas en el análisis, pero no se muestran las veintiuna variables excluidas. Estas variables o ítems se codifican utilizando los siguientes indicadores: las dos primeras letras hace referencia al tipo de tarea que el alumno debe realizar en el ítem (*of*: omisión de fonemas, *af*: adición de fonemas, *if*: identificación de fonemas, *os*: omisión de sílabas, *as*: adición de sílabas), el número que aparece a continuación se corresponde con el número del ítem en la prueba original y los ítems seleccionados son, por orden, los siguientes: 26, 17, 10, 13, 25, 6, 23, 20 y 14. La letra después del número es la posición del fonema o de la sílaba de la palabra-estímulo utilizada en la prueba (*i*= inicial, *f*=final, *m*=medial).

En el cuadro 1 comprobamos que los coeficientes de regresión van asociados a un valor de t significativo, y que los estadísticos de colinealidad, medido a través del coeficiente de tolerancia (T), se aproximan a 1. Este coeficiente nos permite considerar que existe una ausencia de multicolinealidad. Es decir, cada una de las variables predictoras es independiente del resto, con lo cual cada una de ellas aporta información relevante de cara a la validez del modelo predictivo.

Cuadro 1. Variables incluidas en el análisis

<b>Variab</b> les	<b>B</b>	<b>t</b>	<b>T</b>
Constante	2,821	7,695 *	
of26i	2,834	11,497 *	0,804
af17f	2,261	8,390 *	0,733
if10m	1,828	7,116 *	0,738
as13i	1,759	6,415 *	0,693
os25m	2,538	8,694 *	0,844
if6i	2,325	6,409 *	0,772
os23i	2,268	8,954 *	0,775
af20m	1,624	6,112 *	0,794
as14i	1,627	6,166 *	0,692

Notas: B: Coeficientes de regresión no estandarizados; t: Prueba t; T: Estadístico de colinealidad (tolerancia); (\*): Significación de t <0,000.

Fuente: Elaboración propia.

Por otro lado, en el cuadro 2 observamos que con las nueve variables incluidas conseguimos un coeficiente de determinación corregido ( $R^2_c$ ) de 0,922. Esto significa que el 92,2% de la puntuación total de PECO es explicada por estas nueve variables que forman parte de la versión abreviada. Por su parte, y como es fácil de suponer, teniendo en cuenta el elevado valor del coeficiente de determinación corregido, el valor F de ANOVA (216,843) es significativo, por tanto, los coeficientes de regresión obtenidos son significativos para la predicción.

Cuadro 2. Resumen del modelo

R	R <sup>2</sup>	R <sup>2</sup> CORREGIDA	ERROR TÍP. DE	DURBIN-WATSON	ANOVA F
0,962	0,926	0,922	1,700	1,883	216,843 p< ,001

Fuente: Elaboración propia.

Además de la ausencia de colinealidad, otra de las condiciones de la aplicación del análisis de regresión es la comprobación de que la relación entre las variables sea lineal, para ello realizamos un estudio de los residuos tipificados en función de los valores pronosticados para la variable dependiente. Pues bien, podemos comprobar en el gráfico 1 cómo los residuos se distribuyen de manera aleatoria al ajustarse claramente a una nube de puntos sin apreciarse ninguna tendencia. Podemos demostrar que la relación entre las variables es lineal y con igualdad de varianzas (homoscedasticidad), y estas condiciones favorece la precisión de la predicción a partir del modelo de regresión. Otro dato a favor del modelo a partir del estudio de los residuos es la posible existencia de autocorrelación que se obtiene mediante el estadístico de Durbin-Watson (1,883) (cuadro 2), y el hecho de que sea próximo a 2 certifica la inexistencia de autocorrelación de los residuos.

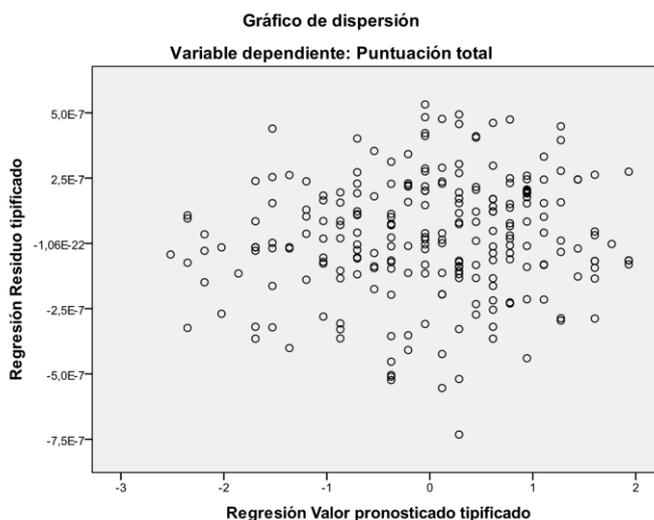


Gráfico 1. Análisis de residuos

Fuente: Elaboración propia.

Por tanto, teniendo en cuenta la evaluación del conocimiento fonológico a partir de los ítems seleccionados por el modelo de regresión, podríamos utilizar los siguientes datos descriptivos y de tipificación e interpretación del nivel de conocimiento fonológico (cuadros 3 y 4). Al ser 9 ítems, la puntuación máxima que se puede obtener es de 9 puntos.

Cuadro 3. Datos descriptivos

MEDIA	DESVIACIÓN TÍPICA	MEDIANA
4,78	2,188	5,00

Fuente: Elaboración propia.

En la tabla de tipificación e interpretación de resultados se muestran las puntuaciones directas correspondientes a los valores cuartiles. Así, en el cuartil 1 ( $Q_1$ ) se sitúan los alumnos cuyo percentil es igual o inferior a 25; es decir, los que obtienen una puntuación directa de 3 puntos o menos, con lo cual consideramos que su nivel de conocimiento fonológico es BAJO, mientras que los que se sitúan en el cuartil 2 y 3 ( $Q_2$  y  $Q_3$ ) serían respectivamente los niveles MEDIO y ALTO.

Cuadro 4. Tipificación e interpretación de los resultados en la aplicación abreviada

	$Q_1$ CENTIL $\leq 25$	$Q_2$ CENTIL ENTRE 26 Y 75	$Q_3$ CENTIL $> 75$
Nivel	Bajo	Medio	Alto
Punt. directa versión abreviada	3 y menos	De 4 a 6	7 y más

Fuente: Elaboración propia.

Por otro lado, presentamos una forma sencilla de obtener los valores pronosticados en la prueba original a partir de la puntuación total en la prueba abreviada. La ecuación de regresión a partir de la variable predictora “Puntuación total Abreviada” (PtA) sería:  $Y'_{pt} = 4,985 + 2,575 * PtA$ . Por ejemplo, a un sujeto que obtiene 5 puntos en la versión abreviada se pronostica una puntuación derivada de la fórmula  $Y'_{pt} = 4,985 + 2,575 * (5) = 17,86$  (por redondeo 18).

En el cuadro 5 presentamos los principales estadísticos del análisis de regresión en la nueva versión abreviada. En este caso, y como es lógico, el valor de coeficiente de determinación ha disminuido en comparación con el del modelo de regresión múltiple, puesto que la predicción se realiza sobre un solo predictor, y por consiguiente el error típico de la estimación aumenta respecto al modelo lineal múltiple, que pasa de 1.700 a 2.264. En cuanto a las características del modelo de regresión simple, también observamos los óptimos valores obtenidos en el coeficiente de determinación, los valores de Durbin-Watson (muy próximo a 2,00) y al elevado valor de F.

Cuadro 5. Resumen del modelo con el predictor PtA

R	R <sup>2</sup>	R <sup>2</sup> CORREGIDA	ERROR TÍP.	DURBIN-WATSON	ANOVA F
0,928	0,861	0,861	2,264	1,986	1479,73 p < ,001

Fuente: Elaboración propia.

Las condiciones para la aplicación del modelo de regresión lineal simple se cumplen en su totalidad. Por un lado, no tiene sentido plantearse la existencia o no de multicolinealidad al existir una sola variable, y por otro lado, el análisis de los residuos (Gráfico 2) demuestran una distribución aleatoria, y por tanto, queda demostrada la linealidad y homoscedasticidad del modelo.

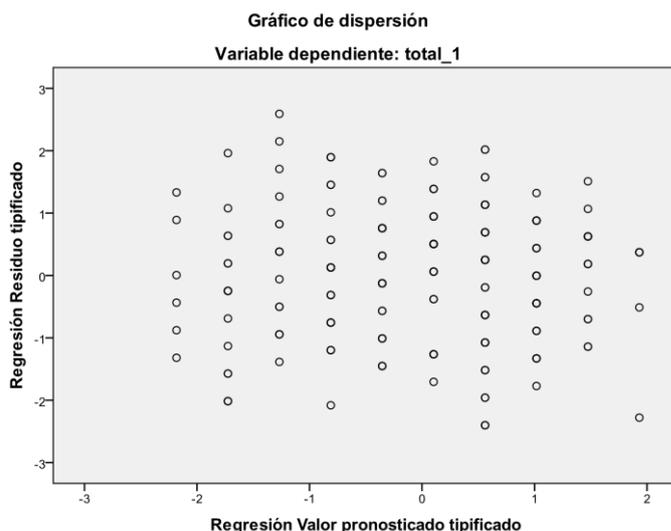


Gráfico 2. Análisis de residuos en la regresión lineal simple  
Fuente: Elaboración propia.

En la gráfica de dispersión (gráfico 3) observamos una relación muy elevada entre la puntuación total en PECO y en la versión abreviada, que cuantitativamente es de 0.928, tal y como aparece en el cuadro 5.

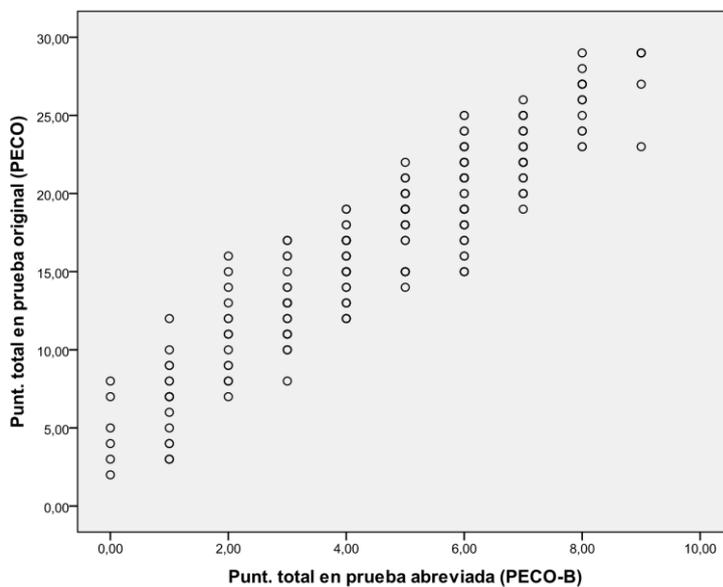


Gráfico 3. Relación entre PECO y versión abreviada  
Fuente: Elaboración propia.

Finalmente, por la sencillez del modelo predictivo y la utilidad para los usuarios, mostramos la tabla de equivalencias pronosticadas entre las PtA en la versión abreviada y en la prueba original PECO (cuadro 6). Además, teniendo en cuenta que cualquier modelo predictivo conlleva un margen de error, obtenemos un error máximo de estimación de 4,44 con un nivel de confianza del 95 %, que sumado y restado al valor

pronosticado en la prueba original tendríamos el intervalo confidencial alrededor del cual se situaría la puntuación pronosticada.

Cuadro 6. Puntuaciones equivalentes pronosticadas en PECO a partir de la versión abreviada.

PUNT. TOTAL ABREVIADA	PUNT. TOTAL PECO	INTERPRETACIÓN CUALITATIVA
9	28	Nivel ALTO
8	26	
7	23	
6	20	Nivel MEDIO
5	18	
4	15	
3	13	Nivel BAJO
2	10	
1	8	

Fuente: Elaboración propia.

De cara a la aplicación de esta versión abreviada, y para ser operativos, recomendamos aplicar esta versión teniendo en cuenta dos criterios. Por un lado, el nivel de dificultad de los ítems expresado por el índice de dificultad; y por otro, por las características de la tarea (*identificación de fonemas – ítems 6, y 10–, adición de sílabas – ítems 13 y 14–, adición de fonemas – ítems 17 y 20–, omisión de sílabas – 23 y 25–, omisión de fonemas – ítem 26–*).

En el cuadro 6 se muestran los índices de dificultad de los nueve ítems que conforman la aplicación abreviada de PECO en el orden propuesto. La interpretación de la dificultad de los ítems se obtiene multiplicando por 100 cada uno de los ítems. Por ejemplo, el ítem “*if6i*” es resuelto correctamente por el 86 % de los sujetos (muy fácil), mientras que el ítem “*af20m*” es resuelto correctamente por el 31 % (difícil).

Cuadro 7. Análisis de dificultad de los ítems y orden de aplicación abreviada

ORDEN DE APLICACIÓN	ÍTEM	ÍNDICE DE DIFICULTAD
1	<i>if6i</i>	0,86
2	<i>if10m</i>	0,55
3	<i>as13i</i>	0,64
4	<i>as14i</i>	0,51
5	<i>af17f</i>	0,65
6	<i>af20m</i>	0,31
7	<i>os23i</i>	0,59
8	<i>os25m</i>	0,21
9	<i>of26i</i>	0,44

Fuente: Elaboración propia.

Para aplicar esta versión abreviada debe disponerse de la prueba original. Siguiendo el orden expuesto en el cuadro 7, se dan al alumno básicamente las siguientes instrucciones:

- *Ítem 6:* De las siguientes palabras (se señalan los dibujos y se dicen las palabras) en qué palabra se oye el sonido /u/ (indio, saco, pie, **uña**, campana).
- *Ítem 10:* De las siguientes palabras (se señalan los dibujos y se dicen las palabras) en qué palabra se oye el sonido /r/ (avión, luna, vela, **barco**, zapato).
- *Ítem 13:* A la ficha blanca la llamaremos /mingo/ (la colocamos primero) mientras que la roja se llama /do/ (la colocamos delante), ¿qué palabra hemos formado?”

(domingo).

- *Ítem 14:* A la ficha blanca la llamaremos /blado/ (la colocamos primero) mientras que la roja sigue llamándose /do/ (la colocamos delante), ¿qué palabra hemos formado? (**doblado**).
- *Ítems 17:* A esta ficha blanca la vamos a llamar /go/ (la colocamos en la mesa) y a esta ficha roja la seguimos llamando /l/, la colocamos detrás de la ficha blanca y decimos: “primero he puesto la ficha blanca que se llama /go/ y después la ficha roja que se llama /l/, ¿qué palabra hemos formado? (**gol**)
- *Ítem 20:* Tenemos tres fichas. La ficha blanca se llamará /pe/ (la colocamos en la mesa). A la ficha amarilla la llamaremos /ado/ (la colocamos en la mesa ligeramente separado de la ficha blanca para dejar hueco a la ficha roja), y a esta ficha roja la llamando /l/, que colocamos en medio de las dos, ¿qué palabra hemos formado? (**pelado**).
- *Ítem 23:* Si a la palabra “saco” le quito el sonido /sa/, ¿cómo sonaría? (**co**).
- *Ítem 25:* Si a la palabra “gusano” le quito el sonido /sa/, (cómo sonaría) (**gu...no**).
- *Ítem 26:* Si a la palabra “foca” le quito el sonido /f/, ¿cómo sonaría? (**oca**).

Nos pareció interesante demostrar además, la validación de la versión abreviada y compararla con la validación en su versión completa. Para ello obtuvimos las siguientes evidencias tanto de validez concurrente como predictiva.

Para la validación concurrente se obtuvo la correlación de Pearson ( $r_{xy}$ ) entre la PtA y los siguientes criterios externos en el mismo período temporal de la medición del conocimiento fonológico: 1) Valoración del maestro-a del “*grado de capacidad del alumno para iniciar la lectoescritura*” según una escala de 0 a 10; 2) Lectura de palabras, y 3) Escritura de palabras.

Para obtener la validez predictiva se utilizaron dos criterios externos (comprensión lectora y escritura de palabras y pseudopalabras) que se correlacionaron con la PtA. Estos dos criterios fueron evaluados al año siguiente de haber aplicado la prueba; esto es, cuando los alumnos se encontraban en el tercer trimestre del primer curso de Educación Primaria. Para evaluar la comprensión lectora se utilizaron las pruebas de “*estructuras sintácticas*” y “*comprensión de textos*” de la batería PROLEC (Cuetos, Rodríguez y Ruano, 1996).

Por su parte, para evaluar la escritura de palabras y pseudopalabras se utilizaron las pruebas de “*escritura de palabras*”, “*escritura de pseudopalabras con sílabas complejas*”, “*palabras de ortografía arbitraria*” y “*palabras con reglas ortográficas*” de la Prueba de evaluación de la escritura en niños de Educación Primaria (Cuetos, Sánchez y Ramos, 1996). La puntuación obtenida en comprensión lectora fue la suma de la puntuación en estructuras sintácticas y comprensión de textos. En el caso de la escritura, se obtuvo una puntuación directa resultante de la suma de las cuatro pruebas de escritura.

En los cuadros 8 y 9 comprobamos que los resultados son bastantes aproximados entre la prueba original y la prueba abreviada. En todos los casos, los coeficientes de correlación son significativos, tanto en la validez concurrente como en la predictiva.

Cuadro 8. Validez concurrente: correlación entre PECO y lectura, escritura y valoración del maestro

	LECTURA PALABRAS (N=116)	ESCRITURA PALABRAS (N=116)	VALORACIÓN MAESTRO-A (N=240)
PECO (abreviada)	0,525 p<0,001	0,588 p<0,001	0,570 p<0,001
PECO (completa)	0,502 p<0,001	0,567 p<0,001	0,616 p<0,001

Fuente: Elaboración propia.

Tabla 9. Validez predictiva: correlación entre PECO y comprensión lectora y escritura de palabras y pseudopalabras

	COMPRESIÓN LECTORA (1° PRIMARIA)	ESCRITURA DE PALABRAS Y PSEUDOPALABRAS (1° PRIMARIA)
PECO (abreviada)	0,374 p<0,001	0,258 p<0,001
PECO (completa)	0,395 p<0,001	0,304 p<0,001

Fuente: Elaboración propia.

En síntesis, a partir de los resultados obtenidos obtuvimos una prueba abreviada y válida del conocimiento fonológico a partir de la prueba original PECO que nos permitirá disponer de un instrumento de evaluación y de investigación de un aspecto tan importante como es el conocimiento fonológico en la adquisición del principio alfabético y en los momentos iniciales del aprendizaje lectoescritor.

## 5. Discusión y conclusiones

Este trabajo aporta una versión abreviada de PECO que constituye una alternativa para la detección de las dificultades en conocimiento fonológico, si tenemos en cuenta que podríamos reducir el tiempo de aplicación a una tercera parte de la aplicación completa. Es decir, la aplicación abreviada se realizaría en 10-12 minutos.

El conjunto de ítems obtenidos en el modelo de regresión constituye una óptima representación de lo que denominamos conocimiento fonológico, puesto que se obtiene una versión no homogénea de ítems (tal y como es considerado teóricamente el constructo “conocimiento fonológico”) y resulta de enorme interés en la evaluación de aquellas variables explicativas de las dificultades de lectura y escritura relacionadas con los procesos fonológicos.

Estamos seguros que, tanto los profesionales con responsabilidad en detectar lo antes posibles las dificultades de lectura y escritura (orientadores, logopedas, psicólogos, pedagogos, etc.) como los investigadores, dispondrán en la versión abreviada de PECO de un adecuado instrumento para evaluar el conocimiento fonológico.

Por otro lado, dada la sencillez de aplicación y la importancia que posee el conocimiento fonológico en los momentos iniciales de la lectura y la escritura, la aplicación abreviada de PECO podría formar parte del conjunto de pruebas que se aplican al finalizar la educación infantil con la finalidad de predecir dificultades lectoescritoras, especialmente de tipo disléxico, y promover las medidas preventivas de refuerzo y apoyo más adecuadas.

## Referencias

- Bravo, L., Villalón, M. y Orellana, E. (2011). La conciencia fonológica y la lectura inicial en niños que ingresan a primer año básico. *Psykhé*, 11(1).
- Byrne, B. y Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, 81(3), 313-321. <https://doi.org/10.1037/0022-0663.81.3.313>
- Carrillo, M. y Sánchez, J. (1991). Segmentación fonológico-silábica y adquisición de la lectura: un estudio empírico. *Comunicación, Lenguaje y Educación*, 9, 109-116.
- Carrillo, M. y Sánchez, J. (1996). *Desarrollo metafonológico y adquisición de la lectura: un programa de entrenamiento*. Madrid: MEC-CIDE.
- Cuetos, F., Sánchez, C. y Ramos, J. L. (1996). Evaluación de los procesos de escritura en niños de educación primaria. *Bordón*, 48(4), 445-465.
- Cuetos, F., Rodríguez, B. y Ruano, E. (1996). *Evaluación de los Procesos Lectores (PROLEC)*. Madrid: TEA ediciones.
- Cutting, L. E. y Denckla, M. B. (2001). The relationship of rapid serial naming and Word reading in normally developing readers: An exploratory model. *Reading and Writing*, 14(7-8), 673-705.
- De Jong, P. F. y Olson, R. K. (2004). Early predictors of letter knowledge. *Journal of Experimental Child Psychology*, 88(3), 254-273. <https://doi.org/10.1016/j.jecp.2004.03.007>
- De Jong, P. F. y Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology*, 91(3), 450-476. <https://doi.org/10.1037/0022-0663.91.3.450>
- Defior, S. y Serrano, F. (2011). Procesos fonológicos explícitos e implícitos, lectura y dislexia. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, 1, 79-94.
- Delgado, M. y Sancho, M. S. (2014). Prerrequisitos para el proceso de aprendizaje de la lectura y la escritura: conciencia fonológica y destrezas orales de la lengua. *Lengua y Habla*, 18, 72-92.
- Domínguez, A. B. y Clemente, M. (1993). ¿Cómo desarrollar secuencialmente el conocimiento fonológico? *Comunicación, Lenguaje y Educación*, 19-20, 171-181.
- Domínguez, A. B. (1992). *La enseñanza de las habilidades de análisis fonológico en el aprendizaje de la lectura y de la escritura. Programas para la Educación Infantil*. Tesis doctoral. Universidad de Salamanca.
- Domínguez, A. B. (1996). Evaluación de los efectos a largo plazo de la enseñanza de habilidades de análisis fonológico en el aprendizaje de la lectura y la escritura. *Infancia y Aprendizaje*, 76, 83-96.
- Feld, V. (2014). Las habilidades fonológicas, su organización neurofisiológica y su aplicación en la educación. *Pensamiento Psicológico*, 12(1), 71-82.
- Gómez, F. R., González, A., Zarabozo, D. y Amano, M. (2010). La velocidad de denominación de letras. *Revista Mexicana de Investigación Educativa*, 15(46), 823-848.
- González, R. M., Cuetos, F., Vilar, J. y Uceira, E. (2015). Efectos de la intervención en conciencia fonológica y velocidad de denominación sobre el aprendizaje de la escritura. *Aula abierta*, 43, 1-8.

- Guzmán, R., Jiménez, J., Ortiz, M., Hernández-Valle, I., Estévez, A., Rodrigo, M., García, E., Díaz, A. y Hernández, S. (2004). Evaluación de la velocidad de nombrar en las dificultades de aprendizaje de la lectura. *Psicothema*, 16, 442-447.
- Jiménez, J. y Ortiz, M. R. (1995). *Conciencia fonológica y aprendizaje de la lectura: teoría, evaluación e intervención*. Madrid: Síntesis.
- Kirby, J. R., Parrila, R. K. y Pfeiffer, S. L. (2003). Naming Speed and Phonological Awareness as Predictors of Reading Development. *Journal of Educational Psychology*, 95(3), 453-464.
- Leal, F. y Suro, J. (2012). Las tareas de conciencia fonológica en preescolar: una revisión de las pruebas empleadas en población hispanohablante. *Revista Mexicana de Investigación Educativa*, 17(54), 729-757.
- López-Escribano, C. (2007). Contribuciones de la neurociencia al diagnóstico y tratamiento educativo de la dislexia del desarrollo. *Revista de Neurología*, 44(3), 173-180.
- Manis, F. R., Seidenberg, M. S. y Doi, L. M. (1999). See Dick RAN: Rapid Naming and the Longitudinal Prediction of Reading Subskills in First and Second Graders. *Scientific Studies of Reading*, 3(2), 129-157. [https://doi.org/10.1207/s1532799xssr0302\\_3](https://doi.org/10.1207/s1532799xssr0302_3)
- Muñoz, C. (2011). Aprendizaje de la lectura y conciencia fonológica: Un enfoque psicolingüístico del proceso de alfabetización inicial. *Psykhe*, 11(1).
- Ortiz, M. R. (1994). *Conciencia fonológica y conciencia del lenguaje escrito en el aprendizaje de la lectura*. Tesis doctoral. Universidad de La Laguna.
- Powell, D., Stainthorp, R., Stuart, M., Garwood, H. y Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology*, 98(1), 46-68. <https://doi.org/10.1016/j.jecp.2007.04.003>
- Ramos, J. L. y Cuadrado, I. (2004). Influence of spoken language on initial acquisition of reading/writing: critical analysis of verbal deficit theory. *Reading Psychology*, 25(3), 149-165.
- Ramos, J. L. y Cuadrado, I. (2006a). *Prueba para la Evaluación del Conocimiento Fonológico – PECO*. Madrid: EOS.
- Ramos, J. L. y Cuadrado, I. (2006b). *Cuaderno para el entrenamiento y refuerzo en conocimiento fonológico*. Madrid: EOS.
- Ramos, J. L. (2000). Entrenamiento en tareas de conocimiento fonológico y su influencia en la adquisición inicial de la lectura y la escritura. *Ciencia Psicológica*, 7, 64-82.
- Ramos, J. L. (2004). Conocimiento fonológico y desarrollo lectoescritor en educación infantil. *Educación XXI*, 7, 169-183.
- Sánchez, E., Rueda, M. I. y Orrantía, J. (1989). Estrategias de intervención para la reeducación de niños con dificultades en el aprendizaje de la lectura y de la escritura. *Infancia y Aprendizaje*, 3-4, 101-111.
- Sebastián, M. E. y Maldonado, A. (1987). *La segmentación de palabras: un prerrequisito del aprendizaje de la lectura*. Boletín I.C.E. Madrid: Universidad Autónoma de Madrid.
- Suárez, P., García, M. y Cuetos, F. (2013). Variables predictoras de la lectura y la escritura en castellano. *Infancia y Aprendizaje*, 36, 77-89.

- Treiman, R. (1991). The role of intrasyllabic units in learning to read. En L. Rieben y C.A. Perfetti (Eds.), *Learning to read: basic research and its implications* (pp. 149-160). Hillsdale: Lawrence Erlbaum Associates.
- Trias, D., Cuadro, A. y Costa, D. (2009). Desarrollo de la conciencia fonémica: evaluación de un programa de intervención. *Ciencias Psicológicas, III*(2), 177-184.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J. y Scanlon, D. M. (2004). Specific Reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2-40. <https://doi.org/10.1046/j.0021-9630.2003.00305.x>
- Wagner, R. K., Torgesen, J. K. y Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*(1), 73-87.
- Yopp, H. (1988). The validity and reliability of phonemic awareness test. *Reading Research Quarterly, 23*, 159-177.

## Breve CV de los autores

### José Luís Ramos Sánchez

Licenciado en Ciencias de la Educación por la UNED y doctor en Pedagogía por la Universidad de Extremadura. Máster en Métodos Actuales de Estadística Aplicada. Profesor Asociado en el área de Métodos de Investigación y Diagnóstico en Educación de la Universidad de Extremadura. Acreditado Profesor Titular de Universidad por la ANECA desde 2009. Desde 1988 es orientador educativo y ha sido Jefe del Servicio de Programas Educativos y Atención a la Diversidad de la Consejería de Educación (Junta de Extremadura). Pertenece al Grupo de Investigación e Innovación Educativa de Extremadura (GRIE) de la UEX. Director de tesis doctorales y trabajos de investigación. Ha participado en proyectos de investigación nacionales y europeos. Autor y coautor de más de cincuenta artículos en revistas científicas nacionales e internacionales y capítulos de libros y libros. Entre sus publicaciones destacan las relacionadas con los métodos de investigación y el análisis de datos en la investigación educativa y social, la evaluación del desarrollo del lenguaje, la conciencia fonológica, la intervención psicopedagógica, evaluación e intervención ante las dificultades de aprendizaje, y en general sobre los procesos de lectura y escritura. ORCID ID: 0000-0002-6636-9383. Email:jlramos@unex.es

### Ana Isabel González Contreras

Licenciada en Psicología por la Universidad de Salamanca y doctora en Psicología por la Universidad de Extremadura. Titulada en psicología clínica por el MEC. Diplomada en profesorado de EGB. Máster Universitario de Investigación en Ciencias Sociales y Jurídicas y Máster en Pedagogía Terapéutica. Profesora asociada del departamento de Psicología y Antropología del área de Psicología Social de la Universidad de Extremadura. Sus líneas de investigación están relacionadas con la convivencia escolar, las dificultades de aprendizaje y el desarrollo socio-familiar. Ha presentado comunicaciones en congresos nacionales e internacionales. En el ámbito privado, ejerce como psicóloga clínica y *Revista Iberoamericana de Evaluación Educativa* 15 educativa. Asimismo, forma parte del turno de

peritaciones judiciales del Colegio Oficial de Psicólogos de Extremadura y de profesionales de la Psicología para la emisión de informes periciales en el ámbito de actuación de tribunales y juzgados en el territorio español. ORCID ID: 0000-0001-8714-6325. Email: [anaisabelgc@unex.es](mailto:anaisabelgc@unex.es)

## Construcción y Validación de Instrumentos para la Evaluación de la Práctica Pedagógica en Educación Básica del Estado de Yucatán

### Construction and Validation of Instruments for the Evaluation of Pedagogical Practice in Elementary Educational Level of the State of Yucatan

Omar Eduardo Kú Hernández \*  
Wilson Jesús Pool Cibrián

Centro de Evaluación Educativa de Estado de Yucatán - México

La evaluación de la práctica pedagógica supone la observación y verificación del quehacer pedagógico del docente, tanto dentro como fuera del aula: fuera del aula rescata los elementos propios de la planeación didáctica, mientras que dentro se refiere al aspecto de la clase como tal y la dinámica de la misma. Aunado a lo anterior, se ha considerado la reflexión del docente sobre su propia práctica pedagógica como un elemento para su mejora y desarrollo. El presente trabajo tiene como propósito la construcción y validación de instrumentos para la evaluación de la práctica pedagógica en las asignaturas académicas de Educación Básica en el Estado de Yucatán. Para esto, se contempló la participación de las figuras del supervisor, director y docente de Educación Básica del Estado de Yucatán, los cuales integraron Comités de trabajo en diferentes momentos del proceso de construcción. Los resultados psicométricos obtenidos y los resultados del grupo de enfoque proporcionan evidencia de instrumentos de medición que cumplen con los requerimientos de validez y confiabilidad para ser implementados en los niveles y modalidades planteadas.

**Palabras clave:** Validación de instrumentos; Instrumentos de evaluación; Evaluación formativa; Evaluación de profesores; Práctica pedagógica; Educación básica; Personal docente.

The evaluation of pedagogical practice involves monitoring and verification of pedagogical work of teachers, both inside and outside the classroom: outside the classroom rescues the elements of educational planning, while in the classroom refers to the aspect of the class as such and dynamics thereof. In addition to this, it is considered the teacher reflection on their own teaching practice as a potentiating element of its improvement and development. This article describes the construction and validation of educational measurement instruments for evaluating teaching practice in academic subjects in elementary educational level in Yucatan. Participated supervisors, directors and teachers of elementary educational level in Yucatan, which integrated working committees at different times of the construction process. Psychometric and focus group results provide evidence of measuring instruments that complied the necessary requirements of validity and reliability to be implemented in the levels and types programmed.

**Keywords:** Instruments validation; Evaluation instruments; Formative evaluation; Teacher evaluation; Teaching practice; Elementary educational level; Teacher.

---

\*Contacto: [omar.ku@yucatan.gob.mx](mailto:omar.ku@yucatan.gob.mx)

issn: 1989-0397

[www.rinace.net/riee/](http://www.rinace.net/riee/)

<https://revistas.uam.es/riee>

Recibido: 14 de julio de 2017

1ª Evaluación: 4 de agosto de 2017

Aceptado: 10 de octubre de 2017

## 1. Introducción

Como parte de los factores que intervienen en el adecuado funcionamiento de los centros escolares, se encuentra la figura del docente. La cual ha sido referenciada por diversos autores como un elemento central en la mejora de la calidad educativa y del logro académico de sus estudiantes (Peña, Fernández y Muñiz, 2009; Rockoff, 2004; Schmelkes, 1996).

En México, la Reforma Integral de la Educación Básica (RIEB) hace énfasis en la actuación de los docentes, como un elemento fundamental para llevar a la práctica la reforma curricular (Ruiz, 2012). Los estudios en la materia han demostrado ampliamente que el profesor es un actor clave en la implementación efectiva de las reformas educativas. Uno de los supuestos principales de la reforma es que los maestros son quienes deben asumir la parte práctica del cambio educativo (Ezpeleta, 2004).

La conciencia de la complejidad de la práctica docente ha llevado a buscar sistemas de indicadores múltiples que combinan pruebas de conocimientos; autorreportes de los docentes, diarios o bitácoras; encuestas que recogen las opiniones de alumnos y padres de familia; observaciones en aula y portafolios y otros acercamientos basados en la revisión de materiales y productos de la docencia. Cada acercamiento tiene puntos fuertes y débiles; ninguno es superior o inferior a otros en forma absoluta, sino que la calidad de la información que proporciona cada uno depende del aspecto a estudiar y del propósito del estudio, de la calidad del instrumento y de la forma en que se aplique y se analicen sus resultados (Martínez Rizo, 2016).

Sin embargo, los esfuerzos del Gobierno Mexicano por hacer obligatoria una evaluación para los maestros de educación básica han provocado una fuerte polémica. Siendo una de las principales preocupaciones, las consecuencias para los docentes en caso de no aprobar su evaluación del desempeño (Martínez Rizo, 2016). El proceso de evaluación del Servicio Profesional Docente (SPD) contempla varias etapas y mecanismos para la verificación de los conocimientos y habilidades de los docentes (SEP, 2016a). Asimismo, proporciona lineamientos a partir de los cuales se establecen los indicadores que deberán cumplirse en el desempeño de los docentes (SEP, 2016b).

Es importante precisar la distinción entre la evaluación del desempeño docente (que es más amplia) y la evaluación de la práctica docente, la cual se restringe a la interacción entre docentes y estudiantes en el ámbito de la clase (Arbesú y Rueda, 2003).

En este trabajo, se especifica sobre la evaluación de la práctica pedagógica, la cual hace referencia al segundo aspecto mencionado. De acuerdo con Moreno (2002), la práctica de los profesores ha recibido varios nombres dependiendo del enfoque sobre el cual se aborde su conceptualización; algunos de estos pueden ser el de pedagógica, docente, educativa, de enseñanza, entre otros.

Hasta hace algunos años, la concepción de la práctica pedagógica aún se limitaba, a una práctica en donde el actor principal era el docente, mientras que el alumno asumía una posición receptiva hacia su aprendizaje (Stone, 1999). Asimismo, dejaba a un lado el contexto en el cual se desarrollaba la actividad educativa y las interacciones que en ella ocurrían. Por su parte, Martínez Rizo (2012) enfatiza el logro de los aprendizajes como la finalidad de la práctica pedagógica y menciona que no es exclusivo del trabajo en el aula,

sino que también se contempla el trabajo fuera del aula que está directamente relacionado con lo que se hace en el aula (p. ej., la planeación de clase). Otro punto de vista señala que tiene que ver con la búsqueda por mejorar su práctica para generar aprendizaje y desarrollo en el alumno y que la actividad docente se encuentra ligada a la realidad contextual del aula (Castro, Peley y Morillo, 2006). De manera similar, Moreno (2002) presenta una perspectiva desde la que se conceptualiza la práctica pedagógica como un proceso de interacción social y educativa en donde intervienen los actores educativos con los lineamientos y políticas institucionales.

Tomando esto en consideración, la evaluación de la práctica pedagógica supone la observación y verificación del quehacer pedagógico del docente, tanto dentro como fuera del aula. Fuera del aula rescata los elementos propios de la planeación didáctica, mientras que dentro del aula se refiere al aspecto de la clase como tal y la dinámica de la misma. Aunado a lo anterior, y desde la visión de los actores involucrados, Bolancé, Cuadrado, Ruiz y Sánchez (2013) mencionan la importancia de considerar la reflexión del docente sobre su propia práctica pedagógica como un elemento para su mejora y desarrollo.

Con respecto a los responsables de realizar esta evaluación, el marco normativo mexicano señala la importancia de los supervisores y directores escolares para la observación y evaluación de la práctica pedagógica, así como de su responsabilidad de éstos para la mejora de la práctica docente (SEP, 2016c). Por otra parte, el marco normativo vigente en el Estado de Yucatán, señala que los supervisores, directores y asesores técnicos pedagógicos tienen la responsabilidad de acompañar la práctica pedagógica y proponer estrategias de mejora (SEGEY, 2011).

Sin embargo, no se cuenta actualmente con herramientas en el Estado que les permitan a estas figuras educativas acercarse a la evaluación de la práctica pedagógica de manera formativa. Si bien existe una iniciativa Federal para la implementación de un método sistemático de observación en el aula (Stallings, 1977, como se citó en SEP, 2015), este método se centra principalmente en verificar el uso y administración del tiempo de la clase, dejando fuera aspectos más relacionados con la calidad educativa como lo refieren otras normativas estatales, nacionales e internacionales (OCDE, 2009; SEGEY, 2011; SEP, 2016b).

Por consiguiente, se plantea este trabajo respondiendo a las necesidades plasmadas en el marco legal de la política educativa en México que hace alusión al seguimiento y atención de los docentes como una piedra angular en el desarrollo educativo. Y también, en función de los lineamientos y estándares plasmados para los docentes por la propia Secretaría de Educación del Gobierno del Estado de Yucatán (SEGEY, 2013-2018). Para esto, se establecen como objetivos: (a) construir instrumentos de evaluación de la práctica docente que permitan retroalimentar su acción pedagógica en el estado de Yucatán; (b) aportar elementos de validez de los instrumentos de evaluación para asegurar que cumplan con características psicométricas adecuadas; (c) verificar la pertinencia de los instrumentos entre la comunidad educativa en el estado de Yucatán a través de la opinión de los actores educativos; y (d) elaborar rúbricas de observación que permitan tener una descripción clara de lo que se espera de cada docente.

## 2. Método

### 2.1. Participantes

Para describir a los participantes en el desarrollo de este proyecto y la aportación que realizó cada uno, se presenta el cuadro 1 en donde se encuentran especificados por actividad.

Cuadro 1. Descripción de las actividades realizadas en el desarrollo del trabajo

COMITÉ/ACTIVIDAD	OBJETIVO	INVOLUCRADOS	CANTIDAD
Comité académico y de especificaciones	Delimitar los contenidos a evaluar y las características de los mecanismos de evaluación	Supervisores, directores y docentes de primaria (general e indígena) y secundaria (en todas sus modalidades)	9 personas: 2 Primaria Indígena 3 Primaria General 4 Secundaria
Comité de elaboración de reactivos	Construir los reactivos en función de las especificaciones proporcionadas por el comité anterior	Colaboradores del CEEY con formación y especialización en el área de ciencias sociales, metodología y educación.	3 personas
Comité de validación de reactivos	Verificar que los reactivos cumplan con las especificaciones establecidas.	Supervisores, directores y docentes de primaria (general e indígena) y secundaria (en todas sus modalidades)	8 personas: 1 Primaria Indígena 3 Primaria General 4 Secundaria
Aplicación Piloto	Utilizar la herramienta en escenarios reales para verificar su pertinencia.	Supervisores de primaria (general e indígena) y secundaria (en todas sus modalidades)	17 personas: 3 Primaria Indígena 3 Primaria General 3 Sec. Estatal 3 Sec. Transferida 3 Sec. Técnica 2 Telesecundaria  Adicional, participaron 88 docentes de primaria y secundaria, quienes fueron los evaluados por sus supervisores y directores.
Elaboración de rúbricas	Diseñar descriptores de los niveles de desempeño por cada reactivo.	Supervisores y directores de primaria (general e indígena) y secundaria (en todas sus modalidades)	14 personas: 2 Primaria Indígena 3 Primaria General 3 Sec. Estatal 3 Sec. Transferida 2 Sec. Técnica 1 Telesecundaria
Validación de rúbricas	Verificar la pertinencia de los descriptores	Supervisores y directores de primaria (general e	32 personas: 6 Primaria Indígena 6 Primaria General

COMITÉ/ACTIVIDAD	OBJETIVO	INVOLUCRADOS	CANTIDAD
	diseñados por el Comité de elaboración de rúbricas	indígena) y secundaria (en todas sus modalidades)	5 Sec. Estatal 5 Sec. Transferida 5 Sec. Técnica 5 Telesecundaria

Fuente: Elaboración propia.

## 2.2. Instrumentos

El procedimiento propuesto en este documento incluye la utilización de tres instrumentos de medición que coadyuvarán a la evaluación de la práctica pedagógica en las asignaturas académicas de Educación Básica:

1. Una guía para la revisión de la planeación didáctica (ver cuadro 2).
2. Una guía de observación en el aula para la verificación de la práctica pedagógica (ver cuadro 3).
3. Una lista de cotejo para la autorreflexión de la práctica pedagógica (ver cuadro 4).

Cuadro 2. Categorías conceptuales del instrumento de Revisión de la Planeación Didáctica

CATEGORÍA	DESCRIPCIÓN	REACTIVOS
Conocimiento del currículo	Se refiere a la forma en la que el docente expresa en la planeación la totalidad de los contenidos del programa de la asignatura, así como los objetivos y la forma en que se pretenden abordar durante el desarrollo de la clase.	4
Estrategias didácticas	Indaga sobre la forma en la que el docente expresa la diversidad de estrategias y recursos que pretenden ser utilizados durante la clase, así como la pertinencia de tales estrategias y recursos para el logro de los aprendizajes.	5
Evaluación	Hace referencia a la forma en la que el docente expresa en su planeación la diversidad y pertinencia de los mecanismos de evaluación que utilizará en su clase para la verificación de los aprendizajes logrados.	4

Fuente: Elaboración propia.

Cuadro 3. Categorías conceptuales de la Guía de Observación en el Aula

CATEGORÍA	DESCRIPCIÓN	REACTIVOS
Conocimiento del currículo	Se refiere al nivel de dominio que el docente demuestra sobre los contenidos de la asignatura, así como la actualidad y vinculación de los contenidos de la asignatura con otros temas relacionados.	4
Estrategias didácticas	Se refiere a las estrategias y actividades que el docente realiza durante su clase para favorecer el aprendizaje de sus estudiantes.	5
Recursos didácticos	Hace referencia a los materiales y recursos de los que se vale el docente durante su clase para favorecer el aprendizaje de sus estudiantes.	4

CATEGORÍA	DESCRIPCIÓN	REACTIVOS
Organización de la clase	Indaga sobre la forma en la que el docente organiza su clase, de tal forma que permita el adecuado abordaje de los contenidos y la realización de las estrategias.	4
Evaluación	Hace referencia a la diversidad y pertinencia de los mecanismos de evaluación utilizados por el docente en diferentes momentos de su clase para la verificación del logro de los aprendizajes.	4

Fuente: Elaboración propia.

Cuadro 4. Categorías conceptuales del instrumento de Autorreflexión de la Práctica Pedagógica

CATEGORÍA	DESCRIPCIÓN	REACTIVOS
Conocimiento de los estudiantes sobre cómo aprenden	Indaga la consideración del propio docente sobre la forma en la que selecciona los contenidos, las estrategias, los recursos y los mecanismos de evaluación de manera pertinente y en función de las características de sus estudiantes.	7
Organización del trabajo educativo e intervención didáctica	Se refiere a la consideración del propio docente sobre la forma en la que organiza su clase para aprovechar los tiempos efectivos, mejorar el aprendizaje de sus estudiantes y promover un ambiente positivo para el aprendizaje.	4
Apoyo a los estudiantes para aprender	Hace referencia a la consideración del propio docente sobre la forma en la que hace uso de las estrategias y recursos para favorecer el aprendizaje de sus estudiantes.	7
Características individuales, culturales y de contexto	Se refiere a la consideración del propio docente sobre su conocimiento acerca de las características del contexto y situación de sus estudiantes, así como de su interacción con la comunidad escolar.	2

Fuente: Elaboración propia.

En el caso de la revisión de la planeación didáctica y la guía de observación, se tratan de escalas estimativas con 4 niveles de desempeño. En el cuadro 5 se describen cada uno de los niveles de desempeño que constituyen una primera guía sobre lo que se considera en cada uno de los niveles. Un segundo elemento para dar claridad a la evaluación son las rúbricas de cada reactivo, las cuales proporcionan una guía descriptiva de lo que se espera en cada uno de los niveles de desempeño por reactivo.

Cuadro 5. Niveles de desempeño para los instrumentos de revisión de la planeación didáctica y la guía de observación en el aula

NIVEL DE DESEMPEÑO	DESCRIPCIÓN
Deficiente	El docente no cumple en lo absoluto o lo hace de forma incorrecta con los aspectos señalados en el enunciado del reactivo.
Insuficiente	El docente cumple de forma parcial con los aspectos señalados en el enunciado del reactivo.
Suficiente	El docente cumple de forma aceptable con los aspectos señalados en el enunciado del reactivo.
Sobresaliente	El docente cumple de manera extraordinaria con los aspectos señalados en el enunciado del reactivo.

Fuente: Elaboración propia.

El instrumento de autorreflexión de la práctica pedagógica es una lista de cotejo con dos opciones de respuesta. En el cuadro 6 se describen cada una de las opciones de respuesta de este instrumento.

Cuadro 6. Niveles de desempeño para el instrumento de autorreflexión de la práctica pedagógica

NIVEL DE DESEMPEÑO	DESCRIPCIÓN
No	Se considera esta opción en caso de que el aspecto señalado en el enunciado NO SE CUMPLA o se encuentre EN PROCESO DE LOGRO
Sí	Se considera esta opción en caso de que se cumpla TOTALMENTE (al 100%) con el aspecto señalado en el enunciado.

Fuente: Elaboración propia.

### 2.3. Procedimiento

Se contemplaron las siguientes etapas para el desarrollo de este trabajo:

1. Revisión de la literatura,
2. Construcción de los instrumentos,
3. Aplicación Piloto,
4. Construcción de rúbricas descriptivas,
5. Resultados y productos finales,
6. Difusión.

En el primer aspecto, se realizó una revisión teórica de varios sistemas de evaluación docente en México y en otras partes del mundo. De igual manera, se hizo una recopilación de instrumentos de medición para la evaluación de la práctica pedagógica. Y finalmente, se extrajeron los documentos propios de la normativa nacional y estatal, que contemplan los lineamientos y mecanismos para la evaluación de la práctica pedagógica.

Para la construcción de los instrumentos, se elaboraron procedimientos de acuerdo con criterios técnicos para el desarrollo y uso de instrumentos de evaluación educativa (INEE, 2014). Para ello, se establecieron cuatro comités: (a) académico, (b) de especificaciones, (c) de elaboración de reactivos y, (d) de validación.

Para cada uno de los comités se utilizaron contenidos y formatos establecidos por el Instituto Nacional para la Evaluación de la Educación (INEE) como fuentes de evidencia del trabajo en cada una de las sesiones. De tal manera, que en cada una de las sesiones de trabajo se generaron productos y se establecieron acuerdos de manera grupal. Los comités académico, de especificaciones y de validación se conformaron por supervisores, directores y docentes de los niveles de primaria y secundaria (públicas). Por su parte, el comité de elaboración de reactivos estuvo conformado por evaluadores educativos, quienes redactaron cada uno de los reactivos siguiendo con las especificaciones señaladas por los comités que los antecedieron.

Después del trabajo con los comités, se procedió a la aplicación piloto de los mismos. Para esto, primero se socializó la información con supervisores de los niveles de primaria y secundaria de la SEGEY y posteriormente, se dio la capacitación para la aplicación piloto de los instrumentos. El tiempo para recolectar los datos fue de 3 semanas, en las cuales cada uno de los supervisores tendría que hacer 5 evaluaciones como mínimo. Con estos datos, se realizaron los análisis psicométricos de los instrumentos.

Los instrumentos fueron entregados por los supervisores en una fecha determinada y con estos actores se realizó un grupo de enfoque con la finalidad de recoger sus opiniones acerca de la aplicación piloto de los instrumentos y verificar la idoneidad del proceso utilizando como preguntas guía: (a) los beneficios que percibieron al utilizar los instrumentos, (b) los retos a los que se enfrentaron al utilizar los instrumentos y, (c) los ajustes y recomendaciones para la mejora de los instrumentos.

Para la construcción de las rúbricas descriptivas se establecieron dos Comités: (a) de elaboración de rúbricas y, (b) de validación de rúbricas. En el primero participaron 14 supervisores de primaria y secundaria, en donde la tarea consistía en desarrollar los descriptores por cada uno de los reactivos de los instrumentos. Para esto, se proporcionó capacitación sobre la metodología para el establecimiento de descriptores de desempeño, se definieron en parejas los descriptores para cada uno de los factores (un factor por pareja), se compartió de manera grupal y finalmente, se revisó de manera cruzada. Para la elaboración de cada nivel de desempeño, se asignaron descriptores para el nivel suficiente, posteriormente para el insuficiente y deficiente, y finalmente para el nivel sobresaliente.

En el caso del segundo Comité, se contó con la participación total de 32 supervisores y directores de primaria y secundaria, en donde el objetivo fue verificar y validar los descriptores de las rúbricas elaboradas por el comité anterior. Este comité, se llevó a cabo en dos sesiones: (a) en la primera se trabajó con la guía de observación en el aula y participaron un total de 20 personas, verificando los instrumentos por equipos en un primer momento, y estableciendo acuerdos en plenaria en un segundo momento; (b) en la segunda sesión se trabajó con la revisión de la planeación didáctica y participaron otras 12 personas, siguiendo con la misma metodología que se siguió con los que verificaron los descriptores de la guía de observación en el aula.

Una vez concluido el trabajo con los actores educativos, se hicieron las modificaciones en función de las observaciones generadas a lo largo del proceso, se generaron los instrumentos en versión electrónica para la generación de reportes automatizados y finalmente, se desarrolló el protocolo de los instrumentos y el manual de aplicación de los mismos, en donde se describe la justificación de los mismos, el proceso de construcción, sus características y las normas de uso.

Seguidamente, el producto final se sometió a revisión con los Directores de los niveles educativos involucrados, Secretarios Técnicos y Coordinadores Pedagógicos de la SEGEY para precisar aspectos de la terminología y normatividad estatal en materia de educación.

Finalmente, se diseñó un proceso de capacitación para los supervisores educativos y directivos de primaria y secundaria de cada una de las 16 regiones educativas en las que está dividido el Estado de Yucatán, con la finalidad de prepararlos para el correcto uso de las herramientas generadas en este proceso.

### **3. Resultados**

Se construyeron tres instrumentos de medición educativa para la evaluación formativa de la práctica pedagógica de asignaturas académicas de Educación Básica del Estado de Yucatán (primaria y secundaria). En los cuadros 7, 8 y 9 se presentan los resultados

descriptivos para los reactivos de la revisión de la planeación didáctica, la guía de observación en el aula y la autorreflexión de la práctica pedagógica, respectivamente.

Cuadro 7. Medias y desviaciones estándar para los reactivos del instrumento de Revisión de la Planeación Didáctica

NO.	REACTIVO	M	D.E.
1	Se encuentran descritos los contenidos temáticos de la asignatura en función de los aprendizajes esperados	2,46	,86
2	Está descrita con claridad la secuencia didáctica de la clase	2,20	,92
3	Están contempladas las competencias y aprendizajes esperados que se pretenden alcanzar	2,42	,86
4	Se observa con claridad el enfoque didáctico de la asignatura	2,23	,91
5	Están descritas estrategias de enseñanza-aprendizaje congruentes con los aprendizajes esperados	2,21	,89
6	Contempla el uso de los recursos tecnológicos con los que cuenta el plantel para el logro de los aprendizajes esperados	1,88	1,15
7	Las actividades planteadas toman en consideración las características de sus estudiantes y de la comunidad escolar para el logro de los aprendizajes	2,18	,98
8	Diseña estrategias de enseñanza-aprendizaje que toman en consideración la diversidad presente en el aula	1,97	1,00
9	Vincula las estrategias y contenidos con las actividades de otras asignaturas y programas del currículo académico del nivel	1,67	1,10
10	Los mecanismos de evaluación descritos toman en consideración la diversidad presente en el aula	1,95	1,02
11	Contempla diversos instrumentos que sean pertinentes para la evaluación de los aprendizajes	1,76	1,07
12	Se encuentran establecidos criterios de evaluación congruentes con las actividades y/o temáticas que se abordarán en la clase	1,88	1,19
13	Se considera en la evaluación las competencias a desarrollar de acuerdo con los aprendizajes esperados	1,75	1,12

Fuente: Elaboración propia.

Cuadro 8. Medias y desviaciones estándar para los reactivos de la Guía de Observación en el Aula

NO.	REACTIVO	M	D.E.
1	Demuestra dominio del tema a través de explicaciones claras y propicia la solución oportuna de las dudas de sus alumnos	2,58	,70
2	El contenido presentado en clase refleja la realidad actual en esa temática	2,54	,67
3	Aborda los contenidos relacionándolos con aspectos significativos para sus estudiantes	2,30	,86
4	Los contenidos presentados son pertinentes para el contexto social de sus estudiantes	2,49	,70
5	Emplea dinámicas y/o actividades lúdicas para promover la participación e interés del estudiante en el proceso de enseñanza-aprendizaje	1,96	,91
6	Favorece la solución de situaciones problemáticas a través de estrategias innovadoras en clase	2,00	,83
7	Promueve la construcción de los aprendizajes en conjunto con sus estudiantes	2,28	,77
8	Emplea estrategias de trabajo colaborativo que promueven la integración y participación de todos sus estudiantes	2,12	,87
9	Utiliza estrategias en diferentes momentos de la clase que favorecen la recuperación, integración y consolidación del conocimiento.	1,95	,89

<b>No.</b>	<b>REACTIVO</b>	<b>M</b>	<b>D.E.</b>
10	Utiliza recursos adecuados para favorecer el logro de los aprendizajes esperados de sus estudiantes	2,22	,71
11	Los recursos empleados por el docente logran centrar el interés de sus estudiantes en la actividad de la clase	2,14	,82
12	Emplea diversos recursos en la dinámica de la clase considerando los distintos estilos de aprendizaje de sus estudiantes	1,84	,89
13	Emplea recursos que son amigables con el medio ambiente	1,89	,96
14	Optimiza el tiempo disponible de la clase para el logro de los aprendizajes esperados	2,25	,83
15	Desarrolla de manera eficiente las actividades de la clase para el logro de los aprendizajes esperados	2,13	,84
16	Favorece un ambiente participativo dentro del aula que facilita el aprendizaje de sus estudiantes	2,34	,80
17	Promueve la creación de un ambiente de convivencia saludable en el aula, basado en el respeto entre todos los actores educativos	2,54	,73
18	Emplea instrumentos que son pertinentes para la evaluación de los aprendizajes	1,79	,85
19	Los mecanismos e instrumentos de evaluación empleados permiten identificar con claridad el nivel de logro alcanzado en cada clase	1,81	,92
20	Los mecanismos e instrumentos de evaluación empleados permiten identificar con claridad los aspectos en los que el estudiante requiere mejorar	1,76	,91
21	Utiliza los resultados de la evaluación para retroalimentar a sus estudiantes de forma escrita y/o verbal.	1,78	,93

Fuente: Elaboración propia.

Cuadro 9. Porcentajes de respuestas para los reactivos del instrumento de Autorreflexión de la Práctica Pedagógica

<b>No.</b>	<b>REACTIVO</b>	<b>NO (%)</b>	<b>SÍ (%)</b>
1	Selecciono los contenidos a desarrollar tomando en consideración los aprendizajes esperados y las características de desarrollo de mis estudiantes	0	100
2	Diseño estrategias didácticas tomando en consideración los propósitos planteados y las características de mis estudiantes	3,7	96,3
3	Empleo estrategias didácticas que sean atractivas y pertinentes para el aprendizaje de mis estudiantes	3,7	96,3
4	Hago uso de los recursos tecnológicos con los que cuento en el centro escolar para facilitar el aprendizaje de mis estudiantes	25,9	74,1
5	Utilizo diversos mecanismos de evaluación acordes con los contenidos y los propósitos a alcanzar por mis estudiantes	12,3	87,7
6	Diseño procesos de evaluación que permiten considerar la diversidad en el aula, haciendo de la evaluación un proceso justo para todos mis estudiantes	11,1	88,9
7	Retroalimentación a mis estudiantes de manera efectiva haciendo hincapié en los resultados del proceso de evaluación, procurando su mejoría en el aprendizaje	8,8	91,3
8	Promuevo situaciones de enseñanza-aprendizaje donde los estudiantes tengan la oportunidad de trabajar de forma colaborativa	1,2	98,8
9	Propicio un ambiente en el aula basado en el respeto y cordialidad para facilitar el aprendizaje de mis estudiantes	1,2	98,8
10	Organizo mi clase de tal forma que se optimice el tiempo disponible para uso exclusivo de actividades académicas	8,8	91,3
11	Establezco normas de comportamiento aceptables en conjunto con mis estudiantes	3,7	96,3

NO.	REACTIVO	No (%)	Sí (%)
12	Utilizo diversos textos de apoyo para aumentar la claridad y comprensión de mis estudiantes en los temas desarrollados en la clase	13,6	86,4
13	Empleo recursos didácticos que favorezcan el involucramiento y la participación de mis estudiantes en las actividades de la clase	8,8	91,3
14	Pongo en práctica estrategias didácticas que promuevan el interés y la participación de todos mis estudiantes	1,3	98,8
15	Investigo estrategias para favorecer el aprendizaje de mis estudiantes	11,4	88,6
16	Utilizo las TIC's como una herramienta para facilitar el aprendizaje de mis estudiantes	19,8	80,2
17	Procuro la participación de otros agentes (padres y docentes) en las actividades para el mejoramiento de los aprendizajes de mis estudiantes	39,5	60,5
18	Trabajo en colaboración con otros profesionales en el enriquecimiento de las actividades que se utilizan para lograr los aprendizajes esperados	35,8	64,2
19	Conozco las características más sobresalientes de mis estudiantes y su contexto sociocultural que podrían influir en su aprendizaje	4,9	95,1
20	Promuevo acciones para interactuar con la comunidad escolar y dar a conocer el avance de mis estudiantes	21,3	78,8

Fuente: Elaboración propia.

#### 4. Validez y confiabilidad de los instrumentos

A partir de la aplicación piloto, se realizó la discriminación de reactivos, el análisis factorial y el análisis de confiabilidad de los instrumentos de la revisión de la planeación y la guía de observación en el aula, dado que son los que utilizan los observadores. En el caso del instrumento de autorreflexión de la práctica pedagógica, únicamente se utilizó la validación por juicio de expertos.

Mediante un procedimiento de grupo superior e inferior, se utilizó una prueba *t de Student* para el análisis de discriminación de reactivos, teniéndose que cada uno de los reactivos de los dos instrumentos resultó ser discriminante (ver cuadros 10 y 11).

Cuadro 10. Análisis de discriminación para los reactivos del instrumento de Revisión de la Planeación Didáctica

REACTIVO	M		D.E.		T	PROB.
	GRUPO INF.	GRUPO SUP.	GRUPO INF.	GRUPO SUP.		
1	1,50	2,91	1,15	0,29	-5,063	,000
2	1,25	2,91	0,79	0,29	-8,886	,000
3	1,65	2,91	1,14	0,29	-4,809	,000
4	1,12	2,86	0,86	0,36	-7,828	,000
5	1,30	2,95	0,98	0,21	-7,402	,000
6	,47	2,64	0,94	0,49	-9,280	,000
7	1,00	2,86	1,00	0,35	-7,722	,000
8	,85	2,77	0,88	0,43	-8,902	,000
9	,50	2,64	0,79	0,58	-9,880	,000
10	,72	2,77	0,83	0,43	-9,529	,000
11	,53	2,77	0,77	0,43	-11,266	,000
12	,35	2,91	0,79	0,29	-12,738	,000
13	,39	2,77	0,70	0,53	-12,295	,000

Fuente: Elaboración propia.

Cuadro 11. Análisis de discriminación para los reactivos de la Guía de Observación en el Aula

REACTIVO	M		D.E.		T	PROB.
	GRUPO INF.	GRUPO SUP.	GRUPO INF.	GRUPO SUP.		
1	1,74	2,96	,87	,20	-5,978	,000
2	1,75	2,92	,72	,28	-6,853	,000
3	1,17	3,00	,79	,00	-9,897	,000
4	1,72	2,92	,83	,28	-5,880	,000
5	,95	2,67	,72	,48	-9,534	,000
6	1,15	2,79	,67	,41	-9,936	,000
7	1,41	2,96	,80	,20	-8,862	,000
8	1,18	2,71	,80	,55	-7,508	,000
9	,84	2,75	,69	,44	-11,023	,000
10	1,67	2,71	,73	,46	-5,618	,000
11	1,36	2,75	,79	,44	-7,258	,000
12	1,00	2,58	,89	,50	-7,436	,000
13	,95	2,79	,89	,51	-8,225	,000
14	1,35	2,92	,81	,28	-8,217	,000
15	1,38	2,83	,86	,38	-7,118	,000
16	1,45	2,96	,80	,20	-8,561	,000
17	1,52	2,96	,68	,20	-9,312	,000
18	,79	2,63	,42	,58	-12,091	,000
19	,61	2,67	,50	,48	-13,449	,000
20	,65	2,58	,49	,58	-11,144	,000
21	,76	2,54	,44	,72	-9,033	,000

Fuente: Elaboración propia.

Posterior a la discriminación de reactivos, se procedió a la realización de un análisis factorial por componentes principales con rotación ortogonal (Varimax). De esta manera, la conformación final de la estructura de cada uno de los instrumentos quedó de la siguiente manera: (a) para la Revisión de la Planeación Didáctica, las tres categorías originales se agruparon en dos factores que explican el 77,7% de la varianza; (b) para la Guía de Observación en el Aula, las cinco categorías originales se agruparon en tres factores que explican el 73,7% de la varianza.

En los cuadros 12 y 13 se describen la conformación de los factores, los reactivos que componen cada factor y las cargas factoriales de cada reactivo de la Revisión de la Planeación Didáctica y la Guía de Observación en el Aula, respectivamente. Asimismo, se agregan las categorías iniciales a las que pertenecía cada uno de los reactivos.

Cuadro 12. Conformación de los factores del instrumento de Revisión de la Planeación Didáctica a partir del Análisis Factorial

FACTOR	REACTIVO	CARGA FACTORIAL	CATEGORÍA INICIAL
	10	,848	Evaluación
	12	,848	Evaluación
	13	,846	Evaluación
	7	,830	Estrategias didácticas
1	8	,824	Estrategias didácticas
	11	,795	Evaluación
	9	,766	Estrategias didácticas
	6	,731	Estrategias didácticas
	5	,687	Estrategias didácticas
2	3	,917	Conocimiento del currículo

FACTOR	REACTIVO	CARGA FACTORIAL	CATEGORÍA INICIAL
	1	,894	Conocimiento del currículo
	2	,779	Conocimiento del currículo
	4	,739	Conocimiento del currículo

Fuente: Elaboración propia.

Cuadro 13. Conformación de los factores de la Guía de Observación en el Aula a partir del Análisis Factorial

FACTOR	REACTIVO	CARGA FACTORIAL	CATEGORÍA INICIAL
1	11	,819	Recursos didácticos
	12	,786	Recursos didácticos
	5	,779	Estrategias didácticas
	13	,727	Recursos didácticos
	10	,658	Recursos didácticos
	7	,591	Estrategias didácticas
	8	,572	Estrategias didácticas
	9	,550	Estrategias didácticas
	6	,553	Estrategias didácticas
2	16	,550	Organización de la clase
	4	,814	Conocimiento del currículo
	3	,805	Conocimiento del currículo
	2	,781	Conocimiento del currículo
	1	,702	Conocimiento del currículo
	17	,645	Organización de la clase
	14	,560	Organización de la clase
	15	,540	Organización de la clase
3	20	,821	Evaluación
	21	,820	Evaluación
	19	,775	Evaluación
	18	,724	Evaluación

Fuente: Elaboración propia.

Una vez obtenida la estructura final de los instrumentos, se realizó un análisis de confiabilidad a través de un análisis de consistencia interna (Alfa de Cronbach) describiéndose en el cuadro 14 los resultados para la Revisión de la Planeación Didáctica, y en el cuadro 15, para la Guía de Observación en el Aula.

Cuadro 14. Índices de consistencia interna obtenidos para el instrumento de Revisión de la Planeación Didáctica

FACTOR	REACTIVO	ÍNDICE DE CORRELACIÓN	ALFA DE CRONBACH POR FACTOR	ALFA DE CRONBACH TOTAL
1	5	,793	,956	,957
	6	,717		
	7	,836		
	8	,860		
	9	,730		
	10	,886		
	11	,860		
	12	,844		
2	13	,847	,920	
	1	,840		
	2	,801		
	3	,832		
	4	,783		

Fuente: Elaboración propia.

Cuadro 15. Índices de consistencia interna obtenidos para la Guía de Observación en el Aula

FACTOR	REACTIVO	ÍNDICE DE CORRELACIÓN	ALFA DE CRONBACH POR FACTOR	ALFA DE CRONBACH TOTAL
1	5	,804	,945	,969
	6	,756		
	7	,761		
	8	,696		
	9	,812		
	10	,703		
	11	,826		
	12	,866		
13	,837			
2	1	,764	,934	
	2	,766		
	3	,836		
	4	,780		
	14	,723		
	15	,725		
	16	,749		
17	,799			
3	18	,847	,942	
	19	,884		
	20	,914		
	21	,806		

Fuente: Elaboración propia.

## 6. Grupo de enfoque

A partir de la aplicación piloto de los instrumentos, se realizó un grupo de enfoque para recoger las opiniones de los actores educativos con respecto a la utilización de los instrumentos. Los resultados se presentan en el cuadro 16 divididos por las fortalezas que señalaron con respecto a los instrumentos y las recomendaciones que hicieron para su mejora.

Cuadro 16. Retroalimentación obtenida por parte de los supervisores ante la aplicación piloto de los instrumentos

FORTALEZAS	RECOMENDACIONES
Instrumentos fáciles de aplicar	Desarrollo de especificaciones para las asignaturas no académicas
Son específicos para la evaluación de la práctica pedagógica	Homologación del proceso en el sistema educativo del Estado, considerando las particularidades de cada uno de los niveles y modalidades
Son relevantes para el contexto	Mantener un uso formativo de esta evaluación
Proporcionan una guía clara de lo que debe evaluarse al entrar al aula	Capacitar a todos los observadores que harán uso de este procedimiento
Permite tener un acercamiento con los docentes para retroalimentar su práctica pedagógica	Proporcionar guías que permitan estandarizar la calificación por cada uno de los observadores

Fuente: Elaboración propia.

### 6.1. Construcción de rúbricas

A partir de esto, se procedió a la elaboración de rúbricas con la finalidad de que los observadores pudieran tener una descripción más clara de lo que podrían esperar en cada

uno de los niveles de desempeño por parte de los docentes. En el cuadro 17 se presenta el ejemplo de un reactivo de la rúbrica.

Cuadro 17. Rúbrica descriptiva correspondiente al reactivo 13 de la Guía de Observación en el Aula

<b>INSTRUMENTO: GUÍA DE OBSERVACIÓN EN EL AULA</b>			
Reactivo 13: Emplea recursos que son amigables con el medio ambiente			
Niveles de Desempeño			
NIVEL 0 (Deficiente)	NIVEL 1 (Insuficiente)	NIVEL 2 (Suficiente)	NIVEL 3 (Sobresaliente)
Omite recursos didácticos o lo hace de forma incorrecta.	Emplea recursos didácticos correctamente, sin que sean amigables con el medio ambiente.	Emplea correctamente recursos didácticos que son amigables con el medio ambiente.	Emplea correctamente recursos didácticos que son amigables con el medio ambiente y promueve la construcción de este tipo de recursos por parte de sus estudiantes.
Especificaciones			
Considerar que los recursos utilizados sean amigables con el medio ambiente a través de la reutilización de materiales o bien, de la utilización de materiales que no generen un impacto ambiental negativo. Los recursos deben tener congruencia con los aprendizajes esperados y favorecer la conciencia ambiental de sus estudiantes. Es recomendable emplear elementos con los que cuenten en el ecosistema (hojas, piedras, frutos, ramas, etc.), sin que esto le provoque algún daño.			

Fuente: Elaboración propia.

## 7. Discusión

Se planteó como propósito de este trabajo la construcción y validación de instrumentos de evaluación que sirvan para los propósitos que dictan las normativas nacional y estatal en términos de la práctica docente. Para esto, se desarrollaron cuatro objetivos a partir de los cuales se discuten los resultados.

Con referencia al primer objetivo, se logró la construcción de los instrumentos para la evaluación formativa de la práctica pedagógica, los cuales están alineados con la normativa tanto estatal como federal y fueron construidos con la participación de los actores educativos que harán uso de esto. Fue un acierto haber considerado como prioritarios estos dos elementos pues como menciona Poissant (2004), sin la participación activa y colectiva de los actores educativos involucrados, la evaluación no tiene más utilidad que servir como mecanismo de control, o peor aún, de simulación administrativa. La participación activa de los docentes significa que éstos deben ser informados de los resultados, tener interés en su explicación, estar motivados para cambiar, saber por dónde cambiar y ver el cambio reflejado en su promoción. A su vez, la participación colectiva significa que los profesores se hacen corresponsables, mediante procedimientos colegiados, de las características y resultados de un proceso que en esencia es de índole social (Ejea, 2007).

En cuanto al segundo objetivo de este estudio, los aportes de validez revelan una conformación lógica de los instrumentos que se obtuvo en primera instancia, a través de la validez de contenido que le confirieron los comités de construcción de instrumentos de

evaluación y una validez de constructo obtenida a partir del análisis factorial de los instrumentos. De igual manera, se observan índices de consistencia interna altos ( $\geq 0.700$ ), lo cual refleja la congruencia entre los reactivos de cada uno de los factores de los instrumentos. Estos resultados concuerdan con otros esfuerzos que se están realizando en el país, en donde se busca elaborar procesos de evaluación por observación (*in situ*) que estén alineados con el marco normativo en educación (Pedroza y Luna, 2017). Sin embargo, aún quedan otros elementos de validez que deben seguir siendo explorados ya que este estudio se limitó al Análisis Factorial Exploratorio (AFE), por lo que sería conveniente obtener otras medidas que permitan un Análisis Factorial Confirmatorio (AFC) y otros modelos de la Teoría de Respuesta al Ítem (TRI) con muestras más numerosas y representativas del sistema educativo estatal.

Cumpliendo con lo expuesto en el tercer objetivo, se tienen comentarios sobre las fortalezas de los instrumentos y sobre las recomendaciones para su mejora. En el primer aspecto, se consideraron como las principales fortalezas su pertinencia y especificidad, así como la posibilidad de generar un acercamiento para la retroalimentación y mejora de la práctica educativa de los docentes. Tal y como se planteó en este trabajo, la finalidad de los instrumentos incluyen un uso exclusivamente formativo que abra los canales de comunicación para la retroalimentación. En ese sentido, hay que destacar la dimensión formativa que puede tener la evaluación de maestros, al identificar deficiencias particulares de unas prácticas y fortalezas de otras, para diseñar actividades de actualización y otras formas de intervención que mejoren el trabajo de los maestros menos experimentados o menos eficaces.

Es claro entonces, que la evaluación por sí misma no es trascendente, si no va unida de alguna manera al mejoramiento. Para que lo último ocurra, parece necesario que, sin dejar de ser objetiva, la evaluación incluya un elemento “amigable” de apoyo y consejo por pares reconocidos y capacitados. Esta concepción de la evaluación es congruente con una visión de la mejora de las escuelas que considera que dicha mejora sólo se podrá conseguir mediante estrategias basadas en la experiencia y el trabajo profesional de los maestros, quienes pueden hacer que las escuelas se vuelvan organizaciones efectivas, centradas en el aprendizaje, y capaces de mejora continua (Martínez Rizo, 2016).

Por otra parte, en cuanto a las recomendaciones de mejora se hace énfasis en proporcionar guías claras que permitan estandarizar los procesos de evaluación por parte de los observadores. En ese sentido, se elaboraron rúbricas descriptivas que permitieron darle solución al cuarto objetivo. Este procedimiento resultó pertinente puesto que se ha reconocido con anterioridad la necesidad de producir e implementar un conjunto coherente de estándares docentes basados en un marco que defina la buena enseñanza en México, marco y estándares que se espera sean revisados y discutidos por los docentes (OCDE, 2010).

Otra de las recomendaciones que se hicieron a partir de este proceso fue la de capacitar a los actores educativos que harán uso de este procedimiento. De acuerdo con la normativa nacional y estatal se refiere al supervisor escolar, director escolar y asesor técnico pedagógico como los actores con mayor relevancia para la realización de esta tarea, lo cual tiene sustento de otras fuentes nacionales e internacionales que enfatizan la importancia de la participación de estos actores en la evaluación y retroalimentación de la práctica docente (Marcano, 2008; Martínez Rizo, 2016; Manzi, González y Sun, 2011; Van

Bruggen, 2010). No obstante, se ha visto que es posible obtener mediciones más válidas y confiables de pares y directivos cuando ellos reciben capacitación para ejercer este rol (Rockoff y Speroni, 2010). Asimismo, una capacitación sólida a los evaluadores y asegurar que éstos posean las características y competencias requeridas, aparecen como factores claves para el éxito de cualquier sistema de evaluación de profesores (Isoré, 2009, como se citó en Manzi et al., 2011).

Por tanto, se contempla en la agenda la capacitación de los observadores como una estrategia para el desarrollo de sus habilidades de observación, a la par de aumentar las probabilidades de realizar una evaluación objetiva que promueva el fortalecimiento de la práctica pedagógica. La relevancia de esta estrategia encuentra sustento en la evidencia de que es crítico para los procesos de evaluación las habilidades y competencias de los evaluadores, incluso con mayor ponderación que la calidad de los instrumentos, pues de los evaluadores depende la correcta utilización de los mismos (Martínez Rizo, 2016; Schleicher, 2013).

Por todo lo anterior, se denota la importancia de esta contribución al sistema educativo estatal, dado que ha permitido la generación de herramientas de evaluación (con elementos de validez y confiabilidad), construidas a partir de los actores educativos involucrados directamente en su uso y dejando como un punto central en la agenda, el fortalecimiento de los evaluados y de los evaluadores.

Asimismo, se pone de manifiesto que la construcción de instrumentos de medición va acompañada de una mejora continua que permita ir agregando otros elementos de validez y confiabilidad a las fuentes de medida y que permitan un acercamiento más pertinente y veraz al fenómeno educativo.

## Financiación

Este trabajo ha sido realizado con las facilidades y recursos otorgados por la Secretaría de Educación del Gobierno del Estado de Yucatán (SEGEY) y del Centro de Evaluación Educativa del Estado de Yucatán (CEEEY).

## Referencias

- Arbesú, M. y Rueda, M. (2003). La evaluación de la docencia desde la perspectiva del propio docente. *Reencuentro*, 36, 56-64.
- Bolancé, J., Cuadrado, F., Ruiz, J. y Sánchez, F. (2013). La autoevaluación de la práctica docente como herramienta para la mejora del proceso de enseñanza y aprendizaje del alumnado. *Avances en Supervisión Educativa*, 18. Recuperado de <https://avances.adide.org/index.php/ase/article/view/155>
- Castro, E., Peley, R. y Morillo, R. (2006). La práctica pedagógica y el desarrollo de estrategias instruccionales desde el enfoque constructivista. *Revista de Ciencias Sociales*, 12 (3), 581-587.
- Ejea, G. (2007). *Sobre prácticas docentes, modelos educativos y evaluación*. División de Ciencias Sociales y Humanidades-Unidad Azcapotzalco-UAM. Recuperado de <http://www.azc.uam.mx/socialesyhumanidades/03/reportes/eco/lec/vlec019.pdf>.
- Ezpeleta, J. (2004). Innovaciones educativas. Reflexiones sobre los contextos en su implementación. *Revista Mexicana de Investigación Educativa*, 9(21), 403-424.

- INEE. (2014). *Criterios técnicos para el desarrollo y uso de instrumentos de evaluación educativa, 2014-2015*. Ciudad de México: INEE.
- Manzi, J., González, R. y Sun, Y. (2011). *La evaluación docente en Chile*. Santiago de Chile: Mide UC.
- Marcano, L. (2008). *Modelo de evaluación del desempeño profesional docente como vía para el desarrollo profesional*. La Habana: Editorial Universitaria.
- Martínez Rizo, F. (2012). Procedimientos para el estudio sobre las prácticas docentes. Revisión de la literatura. *Revista Electrónica de Investigación y Evaluación Educativa*, 18(1), 1-22.
- Martínez Rizo, F. (2016). *La evaluación de docentes de educación básica: Una revisión de la experiencia internacional*. Ciudad de México: INEE.
- Moreno, E. A. (2002). Concepciones de práctica pedagógica. *Revista de la Universidad Pedagógica Nacional*, 16. Recuperada de:  
[http://www.pedagogica.edu.co/storage/folios/articulos/fol16\\_11inve.pdf](http://www.pedagogica.edu.co/storage/folios/articulos/fol16_11inve.pdf)
- OCDE, Organización para la Cooperación y el Desarrollo Económicos. (2009). *Informe TALIS. La creación de entornos eficaces de enseñanza y aprendizaje: Síntesis de los primeros resultados*. Madrid: Santillana.
- OCDE, Organización para la Cooperación y el Desarrollo Económicos. (2010). *Mejorar las escuelas: estrategias para la acción en México. Acuerdo de cooperación México-OCDE para mejorar la calidad de la educación*. Recuperado de  
<https://www.oecd.org/edu/school/46216786.pdf>
- Pedroza, L. H. y Luna, E. (2017). Desarrollo y validación de un instrumento para evaluar la práctica docente en educación preescolar. *Revista Iberoamericana de Evaluación Educativa*, 10(1), 109-129. <https://doi.org/10.15366/riee2017.10.1.006>
- Peña, E., Fernández, R. y Muñiz, J. (2009). Estimación del valor añadido de los centros escolares. *Aula Abierta*, 37(1), 3-18.
- Poissant, H. (2004). La evaluación de la enseñanza universitaria: algunas sugerencias para optimizar su utilización. En M. Rueda (Coord.), *¿Es posible evaluar la docencia en la Universidad? Experiencias de México, Canadá, Francia, España y Brasil* (pp. 85-110). Ciudad de México: ANUIES.
- Rockoff, J. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review, Papers and Proceedings*, 94(2), 247-252.
- Rockoff, J. y Speroni, C. (2010). Subjective and objective evaluations of teacher effectiveness. *American Economic Review*, 100 (2), 261-266. <https://doi.org/10.1016/j.labeco.2011.02.004>
- Ruiz, G. (2012). La Reforma Integral de la Educación Básica en México (RIEB) en la educación primaria: desafíos para la formación docente. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 15(1), 51-60.
- Schleicher, A. (2013). *Effective instruments for teacher evaluation: What education ministers and union leaders think*. Recuperado de [http://www.huffingtonpost.com/andreas-schleicher/effective-instruments-for\\_b\\_2954139.html](http://www.huffingtonpost.com/andreas-schleicher/effective-instruments-for_b_2954139.html)
- Schmelkes, S. (1996). La evaluación de los centros escolares. Presentación en el *Taller sobre Evaluación de Docentes y de Centros Educativos*. Recuperado de <http://www.campus-oei.org/calidad/sylvia.htm>
- SEGEY, Secretaría de Educación del Gobierno del Estado de Yucatán. (2011). *Modelo de Gestión Regional para la Educación Básica del Estado de Yucatán*. Recuperado de:

[http://www.educacion.yucatan.gob.mx/cede/multimedia/material\\_descargable/Modelo\\_de\\_Gestion\\_Regional.pdf](http://www.educacion.yucatan.gob.mx/cede/multimedia/material_descargable/Modelo_de_Gestion_Regional.pdf).

Secretaría de Educación del Gobierno del Estado de Yucatán (SEGEY). (2013-2018). *Programa Sectorial de Educación de Calidad 2013-2018. En Gobierno del Estado de Yucatán* (26 de abril de 2014). Diario Oficial del Gobierno del Estado de Yucatán, 32598. Ciudad de México: Secretaría de Educación Pública.

Secretaría de Educación Pública (SEP). (2015). *Observación de Clase: Herramientas para el supervisor*. Ciudad de México: Secretaría de Educación Pública.

Secretaría de Educación Pública (SEP). (2016a). *Etapas, aspectos, métodos e instrumentos. Proceso de evaluación del desempeño docente y técnico docente. Educación Básica*. Recuperado de [http://servicioprofesionaldocente.sep.gob.mx/2016/ba/EAMI\\_dytd.pdf](http://servicioprofesionaldocente.sep.gob.mx/2016/ba/EAMI_dytd.pdf)

Secretaría de Educación Pública (SEP). (2016b). *Perfil, parámetros e Indicadores para Docentes y Técnicos Docentes en Educación Básica*. Recuperado de [http://servicioprofesionaldocente.sep.gob.mx/2016/ba/PPI/Docente\\_Tecdocente.pdf](http://servicioprofesionaldocente.sep.gob.mx/2016/ba/PPI/Docente_Tecdocente.pdf)

Secretaría de Educación Pública (SEP). (2016c). *Perfil, parámetros e Indicadores para personal con funciones de Dirección y de Supervisión en Educación Básica*. Recuperado de [http://servicioprofesionaldocente.sep.gob.mx/2016/ba/PPI/Direccion\\_Supervision.pdf](http://servicioprofesionaldocente.sep.gob.mx/2016/ba/PPI/Direccion_Supervision.pdf)

Stone, M. (1999). *La enseñanza para la comprensión. Vinculación entre la investigación y la práctica*. Buenos Aires: Paidós.

Van Bruggen, J. C. (2010). *Inspectorates of Education in Europe: Some Comparative Remarks about their Tasks and Work*. Nueva York: Standing International Conference of Inspectorates.

## Breve CV de los autores

### **Omar Eduardo Kú Hernández**

Licenciado en Psicología y Maestro en Investigación Educativa por la Universidad Autónoma de Yucatán. Actualmente trabaja como evaluador educativo en el Centro de Evaluación Educativa del Estado de Yucatán y como consultor educativo y organizacional en CID Mérida. Email: [omar.ku@yucatan.gob.mx](mailto:omar.ku@yucatan.gob.mx)

### **Wilson Jesús Pool Cibrián**

Licenciado en Psicología por la Universidad Marista de Mérida, Maestro en Psicología Aplicada por la Universidad Autónoma de Yucatán y Doctor en Psicología por la Universidad Nacional Autónoma de México. Actualmente se desempeña como jefe del Departamento de Análisis y Difusión de la Información del Centro de Evaluación Educativa del Estado de Yucatán. Email: [wipool77@gmail.com](mailto:wipool77@gmail.com)



## Análisis sobre la Evaluación de la Calidad Educativa en América Latina: Caso Colombia

### Analysis on the Evaluation of Educational Quality in Latin America: Case Colombia

Deivis Robinson Mosquera Albornoza\*

Universidad Católica de Oriente

En el presente texto se pretende reflexionar en torno al concepto de evaluación de la calidad educativa, tratando de entender algunos de los elementos que están implícitos en él, como lo es el sentido de subjetividad y por ende la multiplicidad de interpretaciones que este posee, lo anterior teniendo en cuenta que el hecho de aplicar procesos evaluativos en el campo educativo, se encuentra rodeado de innumerables variables, donde las particularidades del contexto, deben iniciar a tener un papel relevante entorno no solo a los instrumentos que se utilizan, sino también en los resultados de las mismas. De igual forma, es relevante el poder analizar el contexto global y local, con respecto a las razones por las cuales este tipo de pruebas no están contribuyendo a la equidad y la diferenciación que debe existir entre las diferentes formas de aprender de los individuos, el pensar en estandarizar desconoce en gran medida la diversidad, las múltiples condiciones sociales, la cultura, los recursos y los contextos en los que se educan los estudiantes. A su vez, el objetivo de este escrito fue realizar un análisis sobre la situación actual de las evaluaciones estandarizadas en América Latina, tomando como base el caso colombiano, debido al aumento de la participación de los países de la región en el desarrollo de estas; es importante mencionar que lo que se busca es mostrar en primer lugar los aspectos que permitan identificar la importancia de dichas pruebas, al igual que las funciones que estas brindan a los sistemas educativos. Es de anotar que la metodología utilizada fue cualitativa, basada en un enfoque hermenéutico que utiliza como técnica fundamental el análisis documental, dando como resultado un estudio monográfico.

**Palabras claves:** Evaluación estandarizada; Calidad educativa; Política educativa; Medición; Pruebas internacionales; Competencias; Sistemas educativos.

In the present text it is intended to reflect on the concept of evaluation of educational quality, trying to understand some of the elements that are implicit in it, as is the sense of subjectivity and therefore the multiplicity of interpretations that it has, previous taking into account that the fact of applying evaluative processes in the educational field, is surrounded by innumerable variables, where the particularities of the context, must begin to have a relevant role not only to the instruments that are used, but also in the results of the same. Similarly, it is relevant to be able to analyze the global and local context, with respect to the reasons why this type of evidence is not contributing to equity and the differentiation that must exist between the different ways of learning of individuals, the thinking about standardizing largely ignores the diversity, the multiple social conditions, the culture, the resources and the contexts in which students are educated. In turn, the objective of this paper was to analyze the current situation of standardized evaluations in Latin America, based on the Colombian case, due to the increased participation of the countries of the region in the development of these; It is important to mention that what is sought is to show in the first place the aspects that allow to identify the importance of said tests, as well as the functions that these provide to the educational systems. It should be noted that the methodology used was qualitative, based on a hermeneutical approach that used documentary analysis as a fundamental technique, resulting in a monographic study.

**Keywords:** Standardized evaluation; Educational quality; Educational policy; Measurement; International tests; Competences; Educational systems.

---

\*Contacto: [deromoal@gmail.com](mailto:deromoal@gmail.com)

issn: 1989-0397

[www.rinace.net/riee/](http://www.rinace.net/riee/)

<https://revistas.uam.es/riee>

Recibido: 17 de agosto de 2017

1ª Evaluación: 26 de octubre de 2017

Aceptado: 15 de noviembre de 2017

## **1. Introducción**

El presente artículo, pretende mostrar elementos que permitan entender un poco sobre el origen del concepto de evaluación de la calidad educativa, entendiendo que este migró del sector económico, donde se buscaba mejorar no solo los niveles de producción, sino también la calidad de los productos que se fabricaban. Además, se buscará mostrar la evolución que ha tenido dicho concepto a lo largo de la historia, con el fin de poder entender los avances o retrocesos que ha podido tener en el contexto educativo.

Asimismo, se pondrán en evidencia tanto los posibles riesgos que pueden estar implícitos en la realización de dichas evaluaciones, y a su vez las ventajas o potencialidades que trae consigo el hecho de llevar a cabo este tipo de proceso al interior de los sistemas educativos, con el fin de co-ayudar en la búsqueda de consensos que permitan encontrar caminos lógicos y pertinentes sobre la definición de lo que en verdad significa el concepto de evaluación de la calidad educativa.

Del mismo modo, se analizarán algunas implicaciones que tiene la evaluación de la calidad educativa en el desarrollo y fortalecimiento de los sistemas educativos de los países, entendiendo estos como la mejor forma de poder contribuir a que las personas puedan alcanzar un verdadero desarrollo a escala humana, donde al aumentar sus años de escolaridad y poder acceder a sistemas educativos de calidad, les permitan aumentar las posibilidades de mejorar sus condiciones de vida y la de los suyos. De igual forma, se busca evidenciar la situación de la evaluación estandarizada en América Latina, denotando los aspectos en los que ha venido mejorando, pero del mismo modo mostrar cada una de las debilidades que la aplicación de este tipo de evaluaciones trae consigo. Es de anotar que no basta con solo conocer los aspectos antes mencionados, sino que se requiere además conocer los propósitos y el uso que se le da a estos resultados, junto con los criterios necesarios para su divulgación.

Además, la intención será la de realizar un análisis sobre el verdadero desafío que trae consigo el aspecto de la calidad técnica de las ya mencionadas evaluaciones estandarizadas y la importancia de poder constituir unas unidades de evaluación que no solo se encarguen de mostrar resultados para decir si se avanzó o no, sino que además analicen las particularidades de las diversas regiones, con el fin de mostrar análisis contextualizados de las mismas.

Por último, se pretende reflexionar sobre la relación entre los exámenes que se realizan y el concepto de calidad educativa, entendiendo que no debería suponerse, que solo porque una institución educativa no obtiene buenos resultados en una evaluación estandarizada, no significa que no esté realizando buenos ejercicios de calidad educativa, ya que dicho concepto, está cargado por una multiplicidad de elementos que inciden en para que ciertos estudiantes puedan obtener mejores resultados en dichos exámenes, por tanto se hace necesario, no solo denotar dicha relación, sino realizar un análisis a los factores que tienen incidencia en la realización de procesos de calidad educativa.

## 2. Objetivos

### 2.1. Objetivo general

Analizar la realidad de la evaluación de la calidad educativa en América latina: caso Colombia.

### 2.2. Objetivos específicos

- Identificar los principales conceptos teóricos que tienen relación con el proceso de evaluación de la calidad educativa.
- Conocer el papel de tienen las pruebas estandarizadas en el mejoramiento de la calidad educativa.
- Analizar los procesos evaluativos como instrumentos de diseño de políticas públicas pertinentes.

## 3. Evaluación de la calidad educativa

Antes de poder entender lo que representa el concepto de evaluación de la calidad educativa, es preciso desglosar tal concepto e iniciar por hablar de la importancia de la relación entre evaluación y calidad educativa, la cual es entendida por Egido (2005, p. 15), como “conceptos estrechamente relacionados, hasta el punto de que no podría entenderse la una sin la otra”. Lo anterior lleva a entender que los procesos educativos en sí mismos, deben estar permeados en su totalidad por escenarios y actores que propendan por la realización de acciones que permitan contribuir no solo a indicadores entorno a la atención masiva de la población estudiantil, sino que dicha atención que se preste se desarrolle con calidad, tratando de encontrar la denominada pertinencia educativa.

En este sentido es importante mostrar, el aporte realizado por González (2004), el cual afirma que:

*Es poco congruente hablar de calidad sin hablar de evaluación, puesto que tildar una cosa como algo que tiene calidad exige realizar una medida, compararla con un referente ideal y elaborar un juicio sobre la adecuación del objeto o sujeto evaluado al referente utilizado. (p. 132)*

De acuerdo con lo anterior, es entendible el grado de complejidad que trae consigo el termino evaluación en el campo educativo ya que los elementos de juicio y referentes de lo que cada sociedad considera ideal, complican aún más, la idea de lograr un consenso generalizado sobre la forma como debería ser una verdadera educación de calidad; y esto es muy lógico, si se entiende que cada comunidad posee sus particularidades, su acervo cultural y también sus potencialidades, que hace mucho más complejo el pretender estandarizar dichos procesos evaluativos.

Además de lo planteado anteriormente, es importante ver que el concepto de calidad educativa tiende a concebirse como difuso, debido a la subjetividad de los sujetos, los cuales cada uno tiene una idea distinta de lo que puede llegar a significar, lo que lleva a necesitar urgentemente un consenso que permite mínimamente establecer ciertos lineamientos necesarios que posibilite no solo su entendimiento, sino también su efectiva aplicabilidad.

En este sentido, Cano (1998) sostiene que:

*El concepto de calidad es ante todo subjetivo, ya que cada consumidor o usuario tiene una idea distinta de lo que se entiende por él. Sin embargo, todo el mundo está de acuerdo en que hablamos de la calidad cuando vemos cubiertas todas nuestras expectativas, tanto si se trata de un producto o de un servicio. (p. 60)*

La anterior afirmación refuerza el hecho que el fin del concepto de calidad en el campo educativo, deberá ser el poder lograr que los actores de las comunidades educativas puedan lograr un grado de satisfacción tal, que les permita crecer no solo en su aspecto académico, sino también en su crecimiento como persona y por ende que permita el resolver de los diversos problemas que pueden afectar dicho contexto.

No obstante, lo anterior, es relevante el mencionar que el concepto de calidad educativa no es propio del ámbito educativo y que este tuvo sus orígenes en el entorno de los negocios, donde la calidad a la que se referían era la que se veía representada en la producción de diferentes productos, lo cual es algo absurdo cuando se trata de asimilar en el campo educativo, ya que es ilógico tratar de comparar la fabricación de un producto, con la formación de seres humanos. Asimismo, el concepto fue evolucionando en la medida en que los diferentes países fueron logrando ciertos logros a nivel de acceso de la población al servicio educativo, ya que después de lograr que las personas se acercaran a las aulas, el acto seguido estaba orientado a que dicho servicio se lograra prestar no solo en las mejores condiciones, sino también con los mejores estándares de calidad posible. Aquí lo importante en estos momentos, no es el conocer de dónde se originó el concepto, sino entender que los procesos de evaluación de la calidad educativa, no debe ser un proceso unilateral donde los dirigentes toman medidas arbitrarias, sino que sean unas decisiones basadas en los resultados obtenidos, teniendo en cuenta que deben convertirse en decisiones informadas, donde se tome en cuenta las opiniones, particularidades y potencialidades de los diversos actores que intervienen en dicho proceso.

Por otra parte, es relevante saber que, debido a la complejidad del término en mención, existen una serie de elementos que rodean su accionar, como lo son los posibles riesgos que la aplicación del concepto de evaluación de la calidad educativa trae consigo; como lo es la dificultad de lograr que dichas evaluaciones sean objetivas, así también como la pretensión de querer generalizar u homogenizar a las instituciones, cuando lo importante es entender la diversidad de cada institución educativa. De igual forma otro de los riesgos se debe a que muchos de los sistemas evaluativos poseen una estructura preestablecida, que no permiten ningún tipo de cambios o sugerencias, quedándose más en los resultados cuantitativos, y desconociendo en muchas ocasiones, los procesos académicos que se llevan a cabo al interior de las instituciones educativas.

Lo anterior es sustentado por Bolívar (1994) el cual plantea que:

*Se ha logrado crear una expectativa infundada, donde la evaluación no es más que un instrumento retórico que sirve para legitimar el funcionamiento actual de los centros, para justificar las reformas o para utilizarse como mecanismo de producción de verdad, en el sentido en que crea realidades. (p. 19)*

Pero además se debe comprender que no todo entorno a dichos procesos evaluativos es malo, ya que también existen posturas que defienden las potencialidades de dichos procesos, entre las que sostiene que la financiación de la educación debe realizarse de forma racional con el fin de lograr un mayor grado de equidad, de igual forma se considera que este tipo de evaluaciones aporta información de gran utilidad sobre el estado en el que

se encuentran los sistemas y que por ende permite diagnosticar la situación en el que este se encuentra.

#### **4. ¿Por qué las pruebas estandarizadas no miden la calidad educativa?**

Es importante el iniciar esta reflexión sobre la pertinencia y efectividad de las pruebas estandarizadas para medir la calidad educativa, partiendo de aclarar algunos interrogantes como, ¿cuál es el verdadero objetivo de las pruebas estandarizadas?, ¿para qué sirven dichas pruebas?, ¿cómo abordan las pruebas estandarizadas el concepto de calidad educativas?

El poder entender los anteriores interrogantes, ayudará a dilucidar de forma acertada el cuestionamiento inicial, es por esto que se debe iniciar por responder lo que significa una prueba estandarizada, la cual según (Popham, 1999, pp. 2-11) afirma que “una prueba estandarizada es cualquier examen que se administra y califica siguiendo un procedimiento estándar predeterminado. Hay dos tipos principales de pruebas estandarizadas: las pruebas de aptitud y las pruebas de logros”.

De acuerdo a lo anterior es claro que las pruebas estandarizadas lo que buscan es realizar una inferencia acertada sobre los conocimientos que poseen los estudiantes en algunas áreas particulares, lo que lleva a pensar que desde su creación este tipo de pruebas no se compromete a evaluar la calidad educativa, ya que su objetivo y alcance se quedan cortos para lograr dicho fin; y esto se entiende desde la perspectiva que dice que un solo instrumento sería incapaz de valorar de forma acertada los procesos de calidad educativa, debido a los múltiples actores, escenarios, contextos y realidades que intervienen en dicho concepto.

De igual forma es de aclarar que las pruebas estandarizadas lo que buscan es realizar inferencias relativas sobre el estado en que se encuentra un estudiante frente a una destreza en particular, lo que posibilita que estos estudiantes se puedan comparar con otros del país. Y es precisamente en el punto anterior donde se presenta la discrepancia, entendiendo que la manera más adecuada de comparar algo, es estableciendo unos ítems de comparación iguales para todos, y como se dijo anteriormente esto sería casi imposible, debido a la multiplicidad de contextos y la gran variedad de formas de enseñar y aprender que poseen los actores del acto educativo.

Además de los elementos anteriores, es de anotar que los docentes se encuentran inmersos en una gran presión por parte de los gobiernos, donde los establecimientos son valorados de acuerdo con los puntajes que obtengan los estudiantes en las pruebas estandarizadas, lo que hace que en muchos de estos establecimientos educativos se preocupen más por preparar a los estudiantes hacia la presentación de dichas pruebas, que de lograr formar buenos seres humanos con calidad. Lo anterior se agrava aún más, cuando empresas externas ofrecen cursos preparatorios que solo buscan instruir a los estudiantes sobre cómo resolver esas pruebas, preocupándose por la instrucción y no por la formación.

En consecuencia con lo anterior existen algunas razones que ayudan a entender que este tipo de pruebas no pueden evaluar la calidad educativa, donde la poca concordancia entre lo que se evalúa y lo que se enseña juegan un papel fundamental para reforzar esta teoría;

esto se evidencia en la falta de unificación de currículos, los cuales al momento de evaluar los conocimientos a los estudiantes, hace difícil encontrar concordancia entre los contenidos enseñados y los evaluados, y aunque para el caso de Colombia existan unos lineamientos curriculares, donde el Ministerio de Educación Nacional de Colombia (1998) afirma que estos deben orientar la construcción de los planes de estudio, es casi imposible garantizar que las instituciones educativas logren enseñar de forma unificada, debido a que en muchos de los encargados de orientar dicho proceso desconocen el contenido de dichos lineamientos.

Otra de las razones por las cuales estas pruebas no pueden evaluar la calidad educativa, es lo que se denomina causalidad confusa, donde el problema está relacionado con los tres factores que contribuyen al puntaje de los estudiantes en las pruebas estandarizadas de logros; primero lo que se enseña en la escuela, segundo la capacidad intelectual innata del estudiante, y tercero el aprendizaje del estudiante fuera de la escuela; y son precisamente los elementos anteriores los que mayor injerencia tienen al momento de pretender valorar la calidad educativa, porque si los elementos que se evalúan, fueran iguales a los que se enseñan en las escuelas, habría algún grado de correlación en los resultados, pero tal y como se mencionó anteriormente, esto no se cumple, además en este factor tienen que ver no solo lo que se enseña en la escuela, sino también la forma como se enseña en estas, donde la formación y capacitación docente hace que se poseen multiplicidades de metodologías para realizar las respectivas transposiciones didácticas.

Asimismo, otro factor determinante es el relacionado con la capacidad intelectual de los estudiantes, donde es claro mencionar que los niños no nacen con las mismas capacidades, lo que hace incoherente que se pretenda hacer ver como si así fuera; lo más pertinentes es entender la gran diversidad de inteligencias que existen, las cuales no son tenidas en cuenta por las pruebas estandarizadas. Al respecto Gardner (1998) afirma que:

*Todos tenemos siete inteligencias modulares, es decir, cada inteligencia es una competencia autónoma e independiente de las otras; se combina de manera adaptativa para el individuo y la cultura, pero no se influyen entre sí. Las inteligencias son: musical, lógico matemática, espacial, lingüística, cenestésico-corporal, intrapersonal e interpersonal. (p. 51)*

Por tanto, las pruebas estandarizadas si pretenden evaluar la calidad educativa, deben asumir un nuevo enfoque que les permita entender que las condiciones intelectuales que poseen los estudiantes son diferentes entre sí e iniciar a concebir el enfoque de inteligencias múltiples, planteado anteriormente.

En consecuencia, el último factor determinante en la causalidad confusa, es el aprendizaje que se da por fuera de la escuela, es aquí donde en el caso de Colombia, existe una diferenciación entre la educación pública de carácter gratuito y la educación privada, donde esta última es la que posee mejores resultados al momento de presentar las pruebas estandarizadas y esto se explica en gran medida por la posibilidad de realización de actividades extracurriculares de manera frecuente, cosa que es difícil en la educación pública, debido a la falta de recursos para poder acceder a esta herramienta didáctica. Siendo así los estudiantes de la educación pública entrarían a competir en desventaja con aquellos estudiantes que pueden gozar no solo de la posibilidad de acceder a un gran número de actividades extracurriculares, sino también a laboratorios bien dotados, instalaciones de gran nivel y recursos ilimitados.

Por tanto, es inaudito que se crea que una prueba estandarizada podrá medir los procesos de calidad que se llevan a cabo en las instituciones educativas y más aún si se menciona otro elemento diferenciador como lo es el nivel socioeconómico, el cual según cual sea el estrato al que pertenezca el estudiante, va a tener más o menos posibilidades de acceso a herramientas y materiales de calidad que ayudarían a tener un mejor o peor proceso educativo de calidad.

## 5. Complejidad de la medición y evaluación de sistemas educativos

Es importante iniciar este análisis, conociendo que es a lo que se denomina una evaluación estandarizada, y es aquí donde (Ravela, 2008) afirma que:

*Se entiende por evaluación estandarizada en gran escala a aquella que permite producir información comparable acerca de los desempeños de estudiantes pertenecientes a distintos contextos culturales y regionales e, incluso, a distintos países, y que ofrece un panorama de la situación de un país. (pp. 45-63)*

En consecuencia, es entendible que la definición anterior permite aclarar que este tipo de evaluaciones tiene su razón de ser en la cantidad de personas en que se pueda aplicar, para poder obtener cierto tipo de información que permita comparar los resultados de diferentes regiones. Además, es de anotar que estas evaluaciones se enfocan en determinar el cambio que han sufrido ciertos aprendizajes y capacidades que demuestran los estudiantes a lo largo de un año lectivo determinado y a su vez la acumulación de conocimientos que estos puedan adquirir a lo largo de su vida estudiantil.

Asimismo, se puede mencionar que a pesar de las concepciones y críticas que poseen este tipo de evaluaciones, también se trata de poner de manifiesto algunas razones por las cuales es importante que estas se realicen, como una forma de conocer la dinámica de los procesos y los resultados de los sistemas educativo; por tanto las evaluaciones estandarizadas pueden servir como base para diseñar o implementar alguna política educativa y a su vez como un instrumento para conocer el estado de los aprendizajes de los estudiantes en ciertos campos del saber. En este sentido se puede mencionar que la mayoría de este tipo de evaluaciones tal y como están en algunos países, solo ayuda a visualizar los resultados educativos de un conjunto de estudiantes y a entregar información sobre algunas capacidades que alcancen los estudiantes.

No cabe duda entonces, que las evaluaciones estandarizadas bien diseñadas y bien ejecutadas, podrían ser una fuente valiosa tanto de información, como de insumo para que los gobiernos puedan establecer políticas educativas pertinentes y contextualizadas, pero se debe entender que dichas evaluaciones solo ofrecen un indicador parcial dentro de todo el universo llamado calidad educativa, a su vez entender que estas evaluaciones solo son un componente de un sistema de evaluación integra del campo educativo, por tanto hace que estas sean necesarias, pero no suficientes.

Por otra parte, es indudable que si se toca la situación de la evaluación estandarizada en América Latina, se debe mencionar que la participación de los países de la región ha venido aumentando tanto en estudios organizados por la OREAL/UNESCO, como en sistemas de evaluación organizados por el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) y en la participación en pruebas internacionales como

las pruebas PISA, a su vez la destacada participación de algunos países en el constituido grupo iberoamericano de PISA, lo cual es relevante, si se pretende entender en contexto los resultados de dichas pruebas; por tanto, se puede mencionar que estas evaluaciones estandarizadas en algunos aspectos están teniendo avances, como por ejemplo en la transparencia y creciente atención en la difusión y uso de los resultados, de igual forma se está trabajando para que cada vez más se pueda pasar de las actuales pruebas normativas a unas pruebas de criterios que se enfoquen en que es lo que verdaderamente el estudiante sabe y es capaz de hacer; asimismo, se ha venido aumentando la preocupación por iniciar investigaciones que se enfoquen en cada vez más factores que tengan incidencia en el proceso de aprendizaje de los estudiantes, como una manera de avanzar en la constitución de un sistema de evaluación cada vez más integral.

No obstante, a lo anterior, estas evaluaciones poseen una serie de debilidades que es necesario que se analicen, como por ejemplo afirma Ferrer (2006), la confusión que suscita el no saber a ciencia cierta la funcionalidad y propósitos de cada tipo de evaluación, como elemento fundamental para poder poseer un acertado sistema evaluativo, por otra parte se necesita que haya una mayor discusión pública sobre la concurrencia entre lo que se evalúa y lo que se enseña, en pocas palabras se necesita que se puedan mejorar las investigaciones sobre las explicaciones del porque se obtienen ciertos resultados, y no solo eso, sino que a su vez se puedan implementar políticas de estado que posibiliten resolver la problemática identificada con dichas evaluaciones.

En concordancia con lo anterior, Murillo y Román (2010) sostienen que:

*El primer paso que deben asumir los gobiernos, debe ser el de definir cuál debe ser el propósito de su sistema de evaluación y para que se usaran los resultados de las mismas, y que es necesario que se establezca un adecuado equilibrio entre la implementación de pruebas nacionales y la participación en las pruebas internacionales. (pp. 97-120)*

Por tanto, se requiere de un sistema de evaluación que desarrolle exámenes de alta calidad técnica, permeado por principios de transparencia y responsabilidad con los resultados, donde se puedan crear y fortalecer unas unidades de evaluación dotadas de elementos técnicos y de calidad, donde su independencia para analizar y publicar resultados sea su principio fundamental y que además se les brinde los presupuestos necesarios para poder operar de la mejor forma.

## **6. La evaluación como instrumento de política educativa**

Para iniciar con el análisis de la evaluación como instrumento de política pública, es importante entender que el enfoque evaluativo abordado es el concerniente a un tipo de evaluación de competencias, que es el que utilizan las instituciones globalizadoras como la OCDE, y que por ende entran a hacer parte de todo tipo de reformas en las agendas de los diferentes países. No cabe duda, que pareciera que nos encontráramos en un proceso expansionista o colonialista en el ámbito educativo, debido a que las agendas de políticas educativas están más influenciadas por las directrices de organismos internacionales que, por las verdaderas necesidades del contexto, ya que son estas las que realizan ciertas exigencias educativas, articuladas a temas económicos.

Lo anterior, nos lleva a analizar cuidadosamente sobre la necesidad de entender los factores inherentes a la obtención no solo de resultados en evaluaciones estandarizadas,

sino también en los procesos de calidad educativa, al respecto Taba (1993) afirma que la evaluación es un proceso mediante el cual se obtiene información que permite determinar si un determinado programa educativo ha arrojado algún efecto; por tanto se puede decir que la realización de este tipo de evaluaciones es necesaria a fin de obtener información constante y real de algunos aspectos que constituyen la calidad educativa, pero que no se debe supeditar solo al aspecto academicista de los resultados de estas pruebas.

Ahora bien, se debe abordar este tema desde la realidad circundante, como lo es el caso de Colombia, donde el tema evaluativo es asumido por el Instituto Colombiano de Fomento de la Educación Superior (ICFES), el cual desde 1968, ha venido encargándose de estos temas. Desde el año 2002, el Ministerio de Educación Nacional de Colombia (MEN) comenzó a cambiar el enfoque de las evaluaciones nacionales, al pasar de evaluar conocimientos a la evaluación de competencias, el cual asume que “las competencias hacen referencia a los conocimientos y habilidades, así como a las estrategias y rutinas necesarias para aplicarlos, que junto con las emociones y actitudes apropiadas permiten atender demandas y tareas a cabalidad” (Simone y Hersh, 2000, p. 4).

En la actualidad el ICFES, es el encargado de realizar las denominadas pruebas saber, que se realizan en los grados tercero y quinto de primaria, y a su vez en el grado noveno de secundaria, y once de la media vocacional, estas se realizan con el fin de realizar un muestreo del avance que pueden ir teniendo los estudiantes a lo largo de su vida estudiantil y aunque en la actualidad estas pruebas han mejorado considerablemente en la estructura de las preguntas, se hace necesario, la inclusión de elementos contextuales para que los resultados sean acordes a las condiciones sociales de los estudiantes. Asimismo, se realizan las pruebas ECAES con el fin de evaluar los aprendizajes y competencias de los estudiantes que concluyen el ciclo de educación superior.

De acuerdo con lo anterior, se evidencia que el organismo encargado del sistema evaluativo del país adopta un concepto de competencia que difiere con el utilizado por los organismos internacionales como la OCDE (2005), el cual plantea que, una “competencia es algo más que únicamente conocimientos y habilidades. Implica la habilidad para enfrentarse a demandas complejas, haciendo uso y movilizando recursos psicosociales (que incluyen habilidades y actitudes) en un contexto particular”. Lo que se plantea es poder utilizar un concepto un poco más completo de lo que es evaluar por competencias donde no solo se base en conocimiento y habilidades, sino que aborde otros elementos complementarios. En este sentido Vasco (2003), afirma que:

*Conjunto de conocimientos, habilidades, actitudes, comprensiones y disposiciones cognitivas, metacognitivas, socioafectivas, comunicativas y psicomotoras apropiadamente relacionadas entre sí para facilitar el desempeño flexible, eficaz y con sentido de una actividad o de cierto tipo de tareas en contextos relativamente nuevos y retadores.* (pp. 10-14)

Lo anterior, lleva a pensar que el sustento teórico que orienta los sistemas evaluativos de un país debería enriquecerse de otros autores que poseen una visión más integral del tema, para así poder buscar una articulación, entre lo que es enseñar por competencias y evaluar por competencias, porque de lo contrario, siempre existirá la disparidad, entre lo que se enseña y lo que se evalúa.

Sin embargo, la preocupación radica en que, si se utilizan solo las pruebas estandarizadas para medir la calidad educativa, hace que se pierda un poco los procesos de evaluación formativa en las aulas, desconociendo la importancia de esta en la formación de ambiente

educativos y sobre todo con los estudiantes que poseen desempeños bajos. Si se quiere avanzar hacia el enfoque de calidad educativa, se debe generar una reflexión crítica y académica de las políticas educativas, donde puedan intervenir todos los actores educativos, y juntos poder construir un consenso sobre lo que significa calidad educativa a partir de las realidades sociales y culturales de cada contexto. En este sentido, Murcia (1991) plantea que:

*La acción evaluativa supone momentos de reflexión crítica, sobre el estado de un proceso del que se tiene una información sistemáticamente recopilada, a la luz de unos principios y propósitos previamente definidos, con el fin de valorar esa información y tomar decisiones encaminadas a reorientar el proceso; es por eso que para nosotros la acción de evaluación se constituye en un elemento dinamizador fundamental para el estudiante, el maestro, la escuela, la familia y la comunidad. (p. 79)*

## 7. Aspectos metodológicos

Es de anotar, que el alcance del presente estudio fue de tipo descriptivo, debido a que se buscó enunciar de forma ordenada y coherente los elementos que fundamentan el proceso de evaluación de la calidad educativa en Colombia, por tanto, se entiende que el enfoque utilizado fue un enfoque cualitativo, orientado por un modelo epistemológico histórico hermenéutico que busca interpretar y comprender los motivos internos de la acción humana, mediante procesos libres, no estructurados, sino sistematizados, que tienen su fuente en la filosofía humanista.

Asimismo, se utilizaron técnicas de análisis documental entendidas estas como un conjunto de operaciones destinadas a representar el contenido y la forma de un documento para facilitar su consulta o recuperación, o incluso para generar la consolidación de conceptos, donde los instrumentos como las fichas bibliográficas y los cuadros de análisis documental fueron relevantes para poder analizar la realidad de los procesos de evaluación de la calidad educativa en Colombia. En ese orden de ideas, se analizaron documentos e informes, a fin de poder obtener no solo datos confiables, sino también información documental relevante que contribuyera a alcanzar el objetivo propuesto.

Los anteriores instrumentos investigativos, ayudaron no solo a detectar, sino a su vez a seleccionar y clasificar los elementos fundamentales de cada texto a partir de procesos de codificación que ayudaron a establecer las categorías principales del presente estudio, entendiendo que dicho proceso hermenéutico sería esencial para entender no solo las variables que tenían relación con el tema en mención, sino también que posibilitaría entender las relaciones que se entretajan entre ellos.

## 8. Conclusiones

Después de conocer las diversas posturas que rodean el concepto de evaluación de la calidad educativa, se llega a la conclusión de que actualmente es difícil llegar a consensos que permitan unificar criterios sobre su correcta aplicabilidad, en lo que si se está de acuerdo es en la necesidad de que los procesos se puedan evaluar con el fin de corregir los elementos que no estén funcionando de la mejor manera y fortalecer los elementos que estén dando resultados.

Otro elemento importante es que para lograr una debida aplicación se deben considerar todas las singularidades de los sistemas educativos, no con el fin de homogenizar, sino por el contrario, de reconocer particularidades de los diversos contextos, los procesos de evaluación en el campo educativo deberán poseer un alto grado de consenso, donde la participación de los actores que intervienen sea muy activa. Lo anterior lo mencionan Harvey y Green (1993) cuando afirman que en una sociedad democrática debe existir espacio para personas que sostengan diferentes visiones. Se puede decir que evidentemente las pruebas estandarizadas no están diseñadas para evaluar la calidad educativa, pero si poseen facilidad para establecer un dominio relativo de contenidos que logran los estudiantes, y se dice relativo ya que estas pruebas no tienen en cuenta los ítems de mayor o menor contestación, sino los intermedios.

De igual forma se puede decir con certeza que si se pretende evaluar la calidad educativa se deben tener en cuenta una serie de factores determinantes, como lo es encontrar concordancia entre lo que se enseña y se evalúa, el entender que los niños no nacen con las mismas capacidades, lo que hace imposible que un instrumento de evaluación se pueda aplicar de la misma forma y pretendiendo obtener los mismos resultados en todos.

Asimismo los maestros deben estudiar y entender la forma como se diseñan y se llevan a cabo las pruebas estandarizadas, llegando a dilucidar que se pregunta y con qué fin se pregunta a sus estudiantes, lo que ayudará a bajar la presión de categorizar erradamente a las escuelas de acuerdo a dichos puntajes obtenidos, explicando a los demás actores que dichos resultados suministran una mirada engañosa de calidad educativa y poder aportar otra serie de evidencias que muestre que el trabajo de educar no se puede medir con una simple prueba estandarizada, que asume que todos los niños son iguales y que viven y poseen las mismas condiciones socioeconómicas.

El análisis de la situación de las evaluaciones estandarizadas en América Latina trae consigo algunas reflexiones finales entorno a la necesidad de entender que la implementación de este tipo de evaluaciones solo tendrá efectos positivos, si se conciben y emplean como un insumo esencial para la toma de decisiones que lleven al diseño de política pública educativa que se encargue de resolver los problemas que se logren detectar en los resultados que se obtengan, ya que si estos resultados no se conocen, ni se utilizan para dicho fin, no tendrían mayor relevancia, ya que no tiene ningún sentido el solo mostrar resultados como una forma de rendir cuentas, sin realizar un análisis contextual e integral de estas evaluaciones.

De igual forma se requiere de la implementación de instrumentos complementarios que contribuyan a tener una visión integral de los resultados, donde no solo importe si el estudiante aprobó o no, sino que también se preocupen por establecer particularidades culturales y de contextos que son relevantes al momento de implementar sistemas evaluativos pertinentes y de calidad, que posibiliten el utilizar dichos resultados para diseñar políticas públicas que ayude a resolver los problemas que se identifiquen con estas pruebas. Además, los gobiernos se deben comprometer en realizar de forma responsable una divulgación de resultados con todos los actores del proceso educativo, para lograr una mayor retroalimentación del proceso, en pocas palabras es necesario que estas evaluaciones se puedan combinar con estudios cualitativos para lograr una mayor integralidad en dichas pruebas y donde la unidades locales de evaluación posea los recursos técnicos y financieros, ya que evaluar bien, requiere de la mejor y mayor inversión

posible. Es de resaltar que Colombia ha realizado esfuerzos, para avanzar en la consolidación de un sistema de evaluación que permita mejorar en las pruebas estandarizadas, tanto internas como externas, por tal razón ha venido diseñando y poniendo en marcha algunas políticas educativas que, aunque son un avance, siguen siendo insuficientes.

Asimismo, tanto las pruebas nacionales como las internacionales permiten ver el bajo nivel de competencias que tienen los estudiantes colombianos en matemáticas. Según el ICFES (2013), el 44% de los estudiantes colombianos finaliza la educación media sin desarrollar las competencias básicas en esta materia, mientras que según PISA este indicador llega al 70%. Por tal razón, en el año 2015 se creó el Índice Sintético de la Calidad educativa, para medir el progreso en las pruebas SABER de un año a otro, la eficiencia o tasas de repitencia, el desempeño que consiste en el puntaje de las pruebas SABER y el ambiente escolar de todas las instituciones educativas del país. Además, ha iniciado con la implementación de la jornada única en los colegios oficiales en Colombia, la cual pretende aumentar un número de horas que permanecen los estudiantes en los planteles educativos, a fin de fortalecer competencias en lenguaje y matemáticas, tendientes a mejorar los resultados en las pruebas estandarizadas.

Es de anotar que las pruebas ya mencionadas pueden tener diferentes funcionalidades, como la posibilidad que le brinda a las familias el poder elegir a las instituciones que posean mejores resultados, de igual forma, pueden servir para la rendición de cuentas, así como también, para realizar pruebas diagnósticas al interior de los establecimientos educativos y en general para aportar al sistema de información general del funcionamiento educativo y a la investigación educativa.

## Referencias

- Baena, G. (1998). *Técnicas de investigación*. Ciudad de México: Editores Mexicanos Unidos.
- Bolívar, A. (1994). La evaluación de centros: entre el control administrativo y la mejora interna. En J. M. Escudero Muñoz y M. T. González González (Eds.), *Profesores y escuela ¿hacia una reconversión* (pp. 251-282). Madrid: Ediciones Pedagógicas.
- Cano, E. (1998). *Evaluación de la calidad educativa*. Madrid: La Muralla.
- CONEVAL. (2012). *Informe de pobreza y evaluación en el estado de Sonora*. Sonora: Coneval.
- CONEVAL. (2012). *Porcentaje de población en situación de pobreza extrema según entidad federativa*. Sonora: Coneval.
- Delval, J. (1992). *Aprender a aprender. I. El desarrollo de la capacidad de pensar*. Madrid: Alhambra Longman.
- Egido, I. (2005). Reflexiones en torno a la evaluación de la calidad educativa. *Tendencias Pedagógicas*, 10, 17-28.
- Ferrer, G. (2006). *Sistemas de Evaluación de aprendizajes en América Latina: Balance y Desafíos*. Lima: PREAL.
- Gardner, H. (1998). *Inteligencias múltiples*. Barcelona: Paidós.
- González, A. (2004). *Evaluación del clima escolar como factor de calidad*. Madrid: La Muralla.

- Harvey, L. y Green, D. (1993). *Defining quality. Assessment & evaluation in higher education*. Birmingham: QHE.
- ICFES. (2013). *Sistema Nacional de Evaluación Estándarizada de la Educación: alineación del examen SABER 11°*. Bogotá: ICFES.
- INEGI. (2013). *www.inegi.org.mx*. Recuperado de <http://www.inegi.org.mx/sistemas/olap/proyectos/bd/consulta.asp?p=17118&c=27769&s=est>
- Ministerio de Educación Nacional de Colombia. (1998). *Lineamientos Curriculares*. Bogotá: Magisterio.
- Murcia, J. (1991). *Proceso pedagógico y evaluación*. Madrid: Antropos.
- Murillo, F. J. y Román, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación*, 53, 97-120.
- OCDE. (2005). *The definition and selection of key competencies*. París: OCDE.
- Popham, W. J. (1999). ¿Por qué las pruebas estandarizadas no miden la calidad educativa? *Educational Leadership*, 56(6), 2-11.
- Ramos, R. Y. (2001). *Educación integral. Una educación holística para el siglo XXI*. París: Desclée de Brouwer, S.A.
- Ravela, P. (2008). Las evaluaciones educativas que América Latina necesita. *Revista Iberoamericana de Evaluación Educativa*, 1(1), 45-63.
- Simone, D. y Hersh, L. (2000). Definition and selection of key competencies. *The ines compendium*. Recuperado de <https://www.google.es/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwiovOzXgbnYAhUH1xQKHU8fDdUQFggwMAE&url=https%3A%2F%2Fwww.oecd.org%2Fpisa%2F35070367.pdf&usg=AOvVaw0cboijnZiyrDTvn15xdTes>
- Taba, H. (1993). *Elaboración del curriculum*. Buenos Aires: Troquel.
- Vasco, C. (2003). *Estándares básicos de calidad para la educación*. Bogotá: Mimeo.

## Breve CV del autor

### Deivis Robinson Mosquera Albornoz

Soy docente nacido en la Ciudad de Istmina, Departamento del Chocó, de nacionalidad colombiano, he realizado la escuela completa para ser maestro, desde la normal superior, pasando por la licenciatura en educación básica y posteriormente la maestría en educación, que han realizado los aportes necesarios para desempeñarme como maestro en una escuela primaria y en la facultad de ciencias de la educación en la Universidad Católica de Oriente en la Ciudad de Rionegro, Antioquia, donde resido, en la actualidad estoy llevando a cabo mis estudios de doctorado en educación, orientando la tesis en la línea de investigación sobre el fortalecimiento de los procesos investigativos en el aula. ORCID ID: 0000-0003-4472-2104. Email: [deromoal@gmail.com](mailto:deromoal@gmail.com)



# Estilos de Enseñanza y Representaciones sobre Evaluación y Aprendizaje

## Teaching Styles and Representations about Evaluation and Learning

Claudia Alejandra Mazzitelli \*<sup>1,2</sup>

Ana María Guirado <sup>1</sup>

María Julieta Laudadio <sup>1,2</sup>

<sup>1</sup> Instituto de Investigaciones en Educación en las Ciencias Experimentales - Argentina

<sup>2</sup> Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) - Argentina

En este artículo presentamos los resultados obtenidos en una investigación realizada con docentes de Ciencias Naturales con el objetivo de identificar y analizar sus representaciones sociales (RS) sobre evaluación y aprendizaje. Trabajamos a partir de dos técnicas de recolección de datos, un cuestionario con formato de escala Likert en la que los profesores manifestaron su opinión en relación a cada una de las afirmaciones incluidas y dos preguntas abiertas en las que solicitamos que definieran desde su opinión qué es evaluación y qué es aprendizaje. La investigación realizada nos ha permitido no sólo caracterizar cada una de las RS -sobre aprendizaje y evaluación-, sino que hemos estudiado de manera conjunta estas representaciones, pudiendo observar las relaciones entre sus elementos y las vinculaciones con los estilos de enseñanza. Es decir, identificamos las RS de aprendizaje y de evaluación que subyacen a los estilos de enseñanza que guían la práctica docente en el aula. A partir de estos resultados consideramos necesario favorecer en los docentes procesos de reflexión que contribuyan a identificar los supuestos que subyacen a los estilos de enseñanza que guían su práctica en el aula y las implicancias en el aprendizaje de sus alumnos.

**Palabras clave:** Psicología social; Estilos de enseñanza; Evaluación; Aprendizaje; Formación de profesores; Ciencias Naturales.

This article presents the results obtained from a research carried out with teachers of Natural Science with the aim of identifying and analyzing their social representations (SR) about evaluation and learning. Two techniques of data collection, a Likert scale format questionnaire and two open questions were used. In the questionnaire, teachers were asked their opinions on the included statements, while the open questions asked about their definition of evaluation and learning. The results have not only allowed to characterize each SR about evaluation and learning, but also to study together both representations, making relations between them and linking them to the teaching styles. That is, we identified the SR about evaluation and learning that underlie the teaching styles that guide the classroom practice. From these results we believe it is necessary to favor reflection processes which contribute to the identification of the assumptions that underlie the teaching styles in the classroom and the implications on the students learning process.

**Keywords:** Social Psychology; Teaching styles; Evaluation; Learning; Teacher training; Natural Science.

---

\*Contacto: mazzitel@ffha.unsj.edu.ar

issn: 1989-0397

www.rinace.net/riee/

https://revistas.uam.es/riee

Recibido: 5 de julio de 2017

1ª Evaluación: 20 de octubre de 2017

Aceptado: 14 de noviembre de 2017

## 1. Introducción<sup>1</sup>

Promover estudios que ayuden a los docentes a identificar sus representaciones referidas a aspectos educativos y la vinculación con sus prácticas pedagógicas constituye una necesidad actual de la enseñanza. Ello posibilitará contribuir a superar la imitación de modelos y la utilización del conocimiento didáctico de forma indiscriminada (Moral Santaella, 2001).

Analizar las prácticas educativas supone no sólo considerar la interacción en el aula, sino que requiere atender otros aspectos que la conforman como, el pensamiento del profesor, la planificación de la enseñanza y los resultados alcanzados con dichas prácticas (García Cabrero, Loredó Enríquez y Carranza Peña, 2008). Además, se hace necesario promover que los profesores reflexionen sobre las implicancias que una determinada representación del conocimiento, la enseñanza y el aprendizaje tiene en las prácticas en el aula (Laudadio, Mazzitelli y Guirado, 2015). A tal fin integrar una perspectiva psicosocial, que supere modelos intraindividuales, contribuirá a una comprensión más amplia de dichas prácticas.

Por ello, consideramos necesario incorporar, en instancias de formación docente inicial y continua, la identificación de las representaciones sociales (RS) de los profesores como así también la reflexión sobre las implicancias que éstas tienen en sus prácticas áulicas y en el aprendizaje. En este trabajo presentamos los resultados de una investigación con docentes de Ciencias Naturales, desarrollada con el objetivo de favorecer la explicitación y el análisis de sus representaciones vinculadas con el aprendizaje y la evaluación.

## 2. Fundamentación teórica

En este apartado presentamos los aportes desde los cuales hemos procurado establecer las vinculaciones entre los estilos de enseñanza y las representaciones sobre evaluación y aprendizaje, los mismos nos brindan el marco referencial a partir del cual analizamos una realidad tan compleja como la práctica docente.

Partimos de la consideración de que la reflexión docente es dinamizadora de cambios en las prácticas, y constituye un aspecto que empieza a cobrar relevancia en el ámbito de formación tanto inicial como continua. Barrea Andaur (2011) señala que la teoría y la práctica, en la actuación docente, deben ser indisociables para proporcionar al quehacer del aula una base epistemológica que marque diferencias con las formas tradicionales de encarar su tarea. Pacheco Lora (2013) afirma que la sistematización de procesos de reflexión entre profesores, puede favorecer la reelaboración de sus representaciones y lograr una mejor vinculación entre la teoría y la práctica docente.

En tal sentido, en investigaciones que se vienen desarrollando, se observa la importancia de que los docentes conozcan sus RS vinculadas a su accionar docente y reflexionen sobre sus posibles implicancias, ya que estas inciden en sus prácticas en el aula y en la institución

---

<sup>1</sup> Realizado en el marco del proyecto: "La enseñanza y el aprendizaje de la Física y de la Química en la formación docente inicial y continua", Universidad Nacional de San Juan.

escolar (Guirado, Mazzitelli y Olivera, 2013; Guirado, Mazzitelli, Olivera y Quiroga, 2013; Kaplan, 1997, 2003; Mazzitelli, Aguilar, Guirado y Olivera, 2009).

Las RS constituyen un abordaje teórico que se inserta en el ámbito de la Psicología Social. El concepto de RS surge con Moscovici, en la década del 60, con su obra *La psychanalyse, son image et son public*, (1961, 1979). Las RS se refieren a formas o modalidades de conocimiento social, se construyen en los intercambios de la vida cotidiana, constituyen un conocimiento de sentido común y se producen en la intersección entre lo psicológico y lo social (Jodelet, 1986). Son formas de conocimiento de tipo práctico, específicas de un grupo, sirven a la comprensión, la explicación y el dominio de los hechos de la vida diaria. Se trata de un modo de conocimiento actual con una dimensión cognoscitiva y simbólica, están sometidas a una lógica doble: la lógica cognitiva y la lógica social (Abric, 2001). Así, las RS constituyen un objeto construido por y para un grupo social (Castorina, Barrerero y Toscano, 2007).

Algunos de los estudios (Guirado, 2014; Laudadio, Mazzitelli y Guirado, 2015; López y Mota, Rodríguez Pineda y Bonilla Pedroza, 2004) analizan las representaciones de ciencia y aprendizaje y su relación con la práctica de los docentes de Ciencias Naturales, así como su posibilidad de cambio a partir de instancias de formación. Entre sus conclusiones indican que detectaron diferencias entre las concepciones de ciencia y aprendizaje de los profesores y su práctica. En relación con las transformaciones de sus representaciones, parecen circunscribirse al discurso y no se reflejan en la práctica en el aula.

Además, otros estudios realizados sobre las representaciones de los profesores con respecto a la enseñanza y al aprendizaje (Guirado, 2014; Laudadio, Mazzitelli y Guirado, 2015) muestran dos orientaciones de la enseñanza. Por un lado, identifican el predominio de un estilo tradicional, de transmisión del conocimiento centrado en la enseñanza y, por otro lado, un estilo centrado en el aprendizaje. A estos estilos de enseñanza subyacen representaciones diferentes de aprendizaje que van desde un aprendizaje memorístico y reproductivo, considerando un alumno pasivo, hasta un aprendizaje constructivo que valora el papel activo de los estudiantes.

Gargallo López, Suárez Rodríguez, Garfella Esteban y Fernández March (2011), mencionan entre las principales características del estilo centrado en la enseñanza- modelo reproductivo, centrado en el profesor, de transmisión de información- la búsqueda de la reproducción como producto de la enseñanza, el uso que se espera de lo aprendido para aprobar o para el futuro, el conocimiento entendido como algo construido por otros en el cual las opiniones del estudiante no se tienen en cuenta, el profesor responsable de la organización y transformación del conocimiento. En este estilo la relación entre el estudiante y el profesor se valora en grado mínimo y es unidireccional. En cambio, el estilo de enseñanza centrado en el aprendizaje tiene como objetivo el cambio conceptual, en donde el uso que se espera de lo aprendido es para la vida y para interpretar la realidad, la responsabilidad de la organización o transformación del conocimiento es compartida y el conocimiento se entiende como algo construido por los alumnos de modo personal.

Otra de las dimensiones en las que se observa una continua preocupación en relación con las prácticas educativas, son las estrategias y los modelos de la evaluación (Castillo, Hernández, Munevar y Portilla, 2014; Umaña Mata, 2014). Así, la evaluación es un tema que suscita constantes debates y cuestionamientos tanto en lo referido a su concepción, como a su función y modalidades. Actualmente hay consenso respecto a la necesidad de

que la evaluación sea reconceptualizada. Tradicionalmente la evaluación consistía en la obtención de resultados numéricos que apuntaban a señalar cuánto se había aprendido y no el qué, el cómo ni el para qué se habían adquirido los nuevos aprendizajes. Se tomaba en cuenta la información adquirida de manera memorística y no los intereses, inquietudes o capacidades reales de los alumnos (Alcaraz Salariche, 2015).

Torres (2013) señala que al considerar el lugar de la evaluación en las prácticas educativas se observa que se han privilegiado la reproducción y el control del conocimiento de los estudiantes, antes que los procesos de construcción. Este autor afirma que "...la evaluación sigue siendo considerada, mayoritariamente, como un proceso desvinculado de la enseñanza y destinado simplemente a medir, acreditar o certificar resultados de aprendizajes" (p. 287).

Por otra parte, Santos Guerra (2005) indica que la evaluación se puede entender desde dos funciones, vinculadas a dos dimensiones distintas: la evaluación como medición (dimensión tecnológica positivista) y la evaluación como comprensión (dimensión crítica/reflexiva). Desde la dimensión positivista la evaluación se lleva a cabo por pruebas estandarizadas, convirtiéndose en una verificación de aprendizajes y en un medio de control social. Es decir, que el docente –que es quien evalúa– establece los criterios, los aplica e interpreta los resultados. Desde la dimensión crítico-reflexiva, la evaluación se entiende como un proceso y no como un momento final y se realiza a través de diferentes instrumentos y medios. Así, la evaluación se constituye en un permanente proceso reflexivo que facilita la reorientación de los procesos de enseñanza y de aprendizaje. En tal sentido la evaluación le brinda al profesor información sobre el aprendizaje de sus alumnos y sobre su propia práctica educativa.

Desde la perspectiva de la teoría de las RS en relación con la evaluación, Salinas Salazar, Isaza Mesa y Parra Mosquera (2006) identifican dos RS predominantes. Una de las representaciones está referida a la medición y a la tensión-control sobre los alumnos, en su núcleo se encuentra la medición y la calificación como principios organizadores, reguladores y estabilizadores de diversas formas de interacción y asumen la evaluación como alternativa de control en el aula. En este sentido los autores advierten que la medición es entendida como una acción de comparación, lo cual requiere siempre de un modelo para la confrontación que permita decidir, es decir, calificar. Además, la medición y la calificación son ejercidas en forma exclusiva por el docente. Al describir la segunda RS, referida a la comprensión y a la revisión para la autorregulación, señalan que en el núcleo aparecen términos como revisar y rastrear. Estas acciones, poseen rasgos afines con la función formativa y pedagógica de la evaluación ya que rastrear se asocia con la indagación y la búsqueda. En tanto que, revisión se vincula con la detección del error asociado a una fase importante y necesaria en el aprendizaje del estudiante. En esta RS se apunta a un papel activo y determinante del estudiante y a una función diferente del docente que evidencia una predisposición al cambio.

### **3. Metodología**

Llevamos adelante una investigación con un enfoque mixto con predominio de metodología cualitativa. Implementamos técnicas acordes al enfoque teórico metodológico de las RS que nos han permitido acceder al contenido de la representación

e inferir las actitudes relacionadas a las mismas (Abric, 2001; Mazzitelli y Guirado, 2010). Atendiendo al enfoque de la investigación se realizó un muestreo no probabilístico estratégico, este tipo de muestreo no determina a priori el tamaño de la muestra sino que se utiliza el criterio de saturación (Cea D'Ancona, 1999).

En la investigación desarrollada con docentes de Ciencias Naturales, realizamos diferentes actividades de indagación de las RS acerca de aspectos relacionados con la enseñanza, el aprendizaje, la imagen de ciencia y la evaluación. Esta investigación se desarrolla en el marco de un proyecto de formación docente inicial y continua en Ciencias Naturales.

A partir del análisis de los datos obtenidos identificamos, entre otros aspectos, los estilos de enseñanza de los docentes, diferenciando los profesores que poseen un estilo centrado en la enseñanza y los que tienen un estilo centrado en el aprendizaje (Laudadio, Mazzitelli y Guirado, 2015). Los resultados obtenidos en relación con los perfiles acerca de los estilos de enseñanza, concuerdan con las últimas investigaciones (Rocha Chávez, 2012, 2013) que señalan la necesidad de identificar los estilos de los docentes ya que sería más claro de esta manera reconocer su posicionamiento y los supuestos que fundamentan sus prácticas en el aula. Además, consideramos «...conveniente seguir profundizando en la relación entre lo que los docentes dicen y las prácticas en el aula...» (Laudadio, Mazzitelli y Guirado, 2015, p. 20).

Atendiendo a esto, en el presente artículo nos focalizaremos en el análisis de las opiniones de dos docentes de cada uno de los grupos identificados en relación con la evaluación y el aprendizaje. Así, los docentes A y B poseen un estilo centrado en el aprendizaje, mientras que para los docentes C y D predomina un estilo centrado en la enseñanza.

Analizaremos los resultados obtenidos a partir de dos técnicas de recolección de datos implementadas de forma escrita e individual con formato de encuesta:

- ✓ Un cuestionario con formato de escala Likert en la que los profesores manifestaron su opinión en relación a cada una de las afirmaciones incluidas.
- ✓ Dos preguntas abiertas en las que solicitamos que definieran desde su opinión qué es evaluación y qué es aprendizaje.

Cuadro 1. Ítems sobre evaluación

ÍTEMES	CONTENIDO DE LOS ÍTEMES
1	Evalúo no sólo para valorar los resultados del alumno sino para obtener información del proceso de aprendizaje e introducir las mejoras necesarias.
2	El mejor método para evaluar a los alumnos es la prueba escrita.
3	Utilizo como método de evaluación únicamente las pruebas escritas.
4	Complemento la prueba escrita como método de evaluación con otros métodos de evaluación formativa/continua (ej. Trabajos prácticos, pequeñas investigaciones, informes, etc.)
5	La evaluación debe limitarse a la valoración de los conocimientos adquiridos.
6	Uso procedimientos de evaluación formativa/continua (p.ej. preguntas de clase, trabajos, informes, pruebas, ensayos, etc.) revisando y devolviendo corregidos a los alumnos los trabajos escritos con instrucciones para la mejora.

Fuente: Elaboración propia.

En los cuadros 1 y 2 presentamos los ítems incluidos en el cuestionario, agrupados según el objeto de la representación al que corresponde, y a los que los docentes debían responder

seleccionando entre las siguientes opciones: 1=Muy de acuerdo; 2=De acuerdo; 3=Indiferente; 4= En desacuerdo; 5= Muy en desacuerdo.

Cuadro 2. Ítems sobre aprendizaje

ÍTEMS	CONTENIDO DE LOS ÍTEMS
1	Basta con que los alumnos aprendan y comprendan los contenidos científicos fundamentales de la disciplina, no necesitan ir más allá en su formación.
2	Aprender es incrementar los conocimientos disponibles.
3	El conocimiento debe ser construido por los estudiantes con ayuda del profesor
4	El contacto con la realidad y el trabajo en el laboratorio son imprescindibles para el aprendizaje de las Ciencias.
5	Aprender es construir personalmente significados.
6	Los conocimientos que mis estudiantes adquieren les sirven para interpretar la realidad en que están inmersos, no sólo para aprobar la materia.
7	El papel básico de los alumnos en clase es estar atentos y tomar bien los apuntes.
8	Las ideas espontáneas de los alumnos deben ser el punto de partida para el aprendizaje de los contenidos científicos.
9	Los alumnos están más capacitados para comprender un contenido si lo pueden relacionar con conocimientos previos que ya poseen.
10	En general, los alumnos son más o menos hábiles según las capacidades innatas que posean.

Fuente: Elaboración propia.

#### 4. Análisis de resultados

A partir de las respuestas dadas al cuestionario graficamos perfiles actitudinales. En estos gráficos se encuentran en el eje de las ordenadas las opciones para las respuestas, mientras que en el eje de las abscisas se ubican los ítems de la escala likert correspondiente.

En primer lugar analizaremos los perfiles para cada docente comparándolos con las respuestas a la pregunta abierta sobre evaluación y aprendizaje (figuras de 1 a la 8). Posteriormente presentaremos los perfiles actitudinales referidos a evaluación y aprendizaje comparando los cuatro docentes seleccionados para su análisis (figuras 9 y 10). De esta manera pretendemos identificar las diferencias individuales y los aspectos comunes al grupo.

A continuación analizamos las respuestas a la escala likert y la definición de evaluación del docente A.

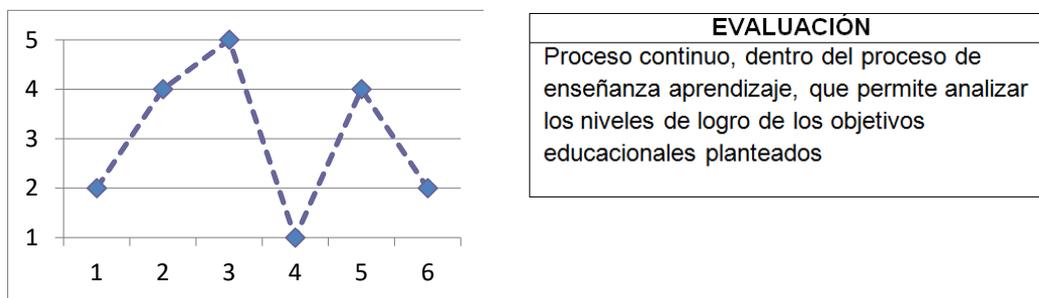


Figura 1. Perfil actitudinal y definición de evaluación del docente A

Fuente: Elaboración propia.

Como podemos observar en la figura 1, el docente A presta acuerdo a los ítems 4 (muy de acuerdo) y a los ítems 1 y 6 (de acuerdo) así este docente expresa que considera la evaluación como formativa/continua y que evalúa para obtener información del proceso de aprendizaje e introducir mejoras. Sus desacuerdos son para los ítems 2 y 5 (en desacuerdo) y 3 (muy en desacuerdo), de este modo el docente considera que la prueba escrita no es el único y mejor método de evaluación y que no sólo evalúa conocimientos adquiridos. Considerando estas opiniones y la definición dada por este profesor, se destaca una representación de la evaluación como parte de los procesos de enseñanza y de aprendizaje y que favorece la reflexión y modificación de estos últimos ya que el docente acuerda con que «permite analizar», lo cual indicaría una atención puesta en los procesos más que en los resultados. Además, la evaluación sería parte del proceso de aprendizaje y de enseñanza ya que se remite también a los objetivos educativos planteados. Así, la representación de evaluación de este docente se vincularía con un estilo de enseñanza centrado en el aprendizaje coherente con la clasificación que ya habíamos identificado en relación con los estilos de enseñanza (Laudadio, Mazzitelli y Guirado, 2015).

En la figura 2 presentamos las respuestas a la encuesta y la definición de aprendizaje del docente A.

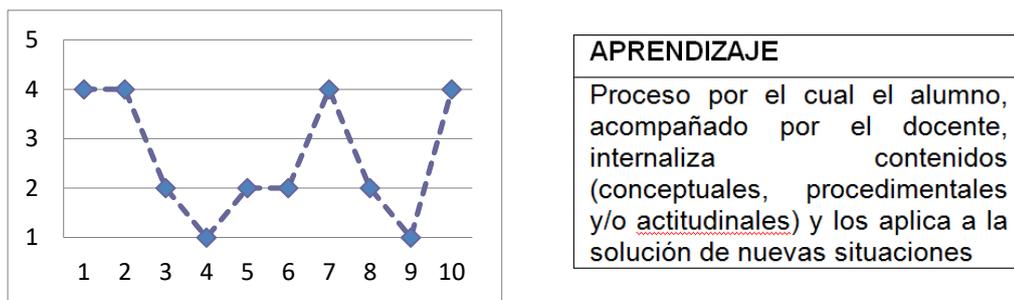


Figura 2. Perfil actitudinal y definición de aprendizaje del docente A  
Fuente: Elaboración propia.

A partir de este perfil (figura 2) podemos observar que el docente no acuerda con los ítems 1, 2, 7 y 10, dichos ítems están relacionados con un alumno que acumula conocimientos, que tiene un rol pasivo frente a los contenidos y que su aprendizaje depende de capacidades innatas. Estos desacuerdos son coherentes con los acuerdos del docente con los ítems 4 y 9 (muy de acuerdo) que hacen referencia al trabajo en laboratorio y la consideración de las ideas previas de los alumnos para el aprendizaje de las ciencias. Además, acuerda con los ítems que hacen referencia a la construcción del conocimiento por parte de los alumnos a partir de sus saberes y la posibilidad de, a partir de ello, interpretar la realidad (ítems 3, 5, 6 y 8). Esto último muestra correspondencia con la definición dada por el docente acerca del aprendizaje, en donde considera al mismo como un proceso, otorga importancia al acompañamiento docente y a la posibilidad de aplicación de los conocimientos a la solución de situaciones nuevas.

Así, teniendo en cuenta las respuestas dadas a la encuesta y la definición de aprendizaje, esta representación del docente A también se vincula con un estilo de enseñanza centrado en el aprendizaje.

En segundo lugar presentamos los resultados a partir de las respuestas del docente B.

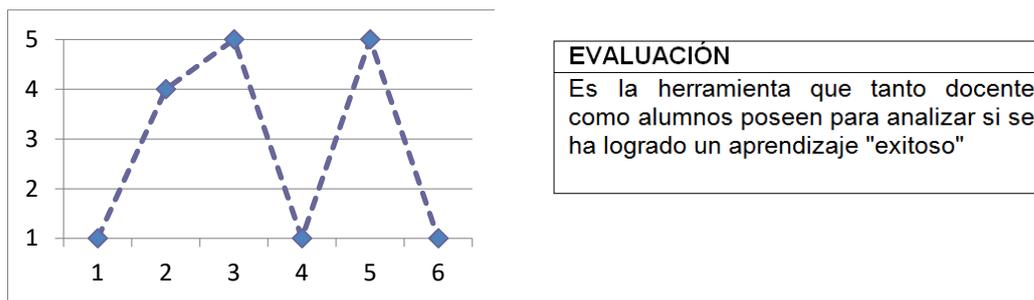


Figura 3. Perfil actitudinal y definición de evaluación del docente B

Fuente: Elaboración propia.

A partir del perfil podemos observar que el docente manifiesta estar muy de acuerdo con los ítems 1, 4 y 6 mostrando de este modo que la evaluación le permite obtener información del proceso de aprendizaje para obtener mejoras y que usa diversos métodos de evaluación tendiendo a una evaluación formativa continua. Manifiesta su desacuerdo con los ítems 2, 3 y 5, que hacen referencia al uso de la prueba escrita como el mejor y único método y a la evaluación como valoración únicamente de conocimientos adquiridos. Por otra parte, en la definición este docente hace referencia a la evaluación como herramienta siendo realizada tanto por docentes como por alumnos con el objetivo de reflexionar sobre el aprendizaje y favorecerlo.

Así, la representación de este docente, similar a la del docente A, se vincularía con un estilo de enseñanza centrado en el aprendizaje.

A continuación analizamos los resultados sobre aprendizaje para el docente B.

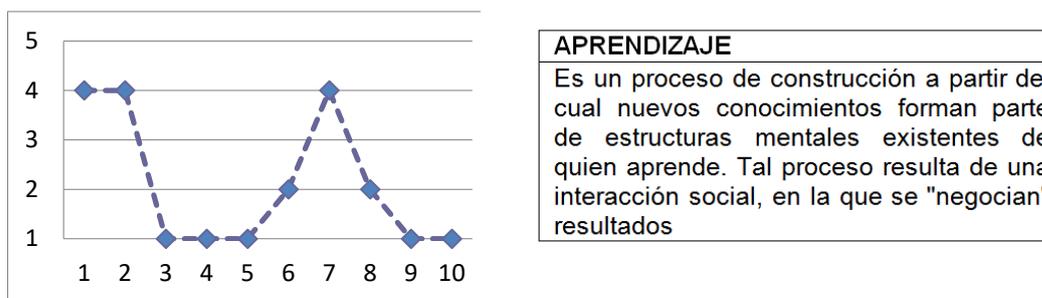


Figura 4. Perfil actitudinal y definición de aprendizaje del docente B

Fuente: Elaboración propia.

En el perfil de aprendizaje del docente B (figura 4) podemos observar que manifiesta estar muy de acuerdo con los ítems 3, 4, 5, 9 y 10 y de acuerdo con los ítems 6 y 8. Así, el docente considera que los estudiantes construyen el conocimiento con ayuda del docente, valoriza las prácticas de laboratorio, la posibilidad de interpretar la realidad desde los conocimientos adquiridos y las ideas previas de los estudiantes como punto de partida. Por otra parte no presta acuerdo a los ítems: 1, 2 y 7 mostrando su desacuerdo con un alumno pasivo y que acumule conocimientos. Respecto de su definición, considera el

aprendizaje como un proceso de construcción teniendo en cuenta los conocimientos existentes, los nuevos conocimientos y la interacción social.

De este análisis se desprenden dos cuestiones una que la representación de este docente se vincularía con un estilo de enseñanza centrado en el aprendizaje y otra, que al igual que con el perfil de evaluación, el perfil de aprendizaje es similar entre el docente A y B.

En tercer lugar presentamos los resultados para la encuesta y las definiciones dadas por el docente C.

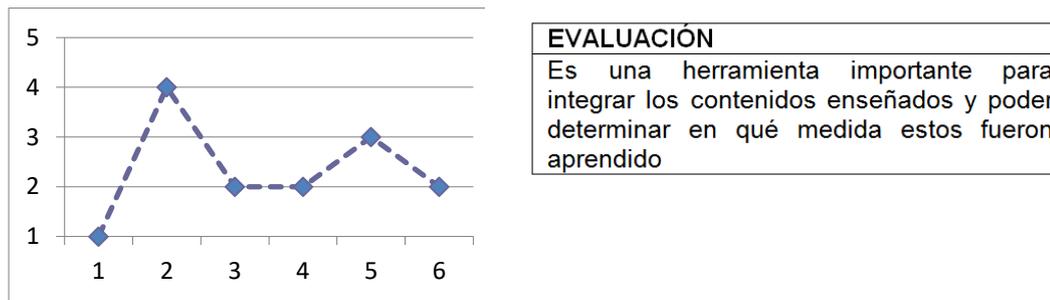


Figura 5. Perfil actitudinal y definición de evaluación del docente C  
Fuente: Elaboración propia.

Al analizar este perfil (figura 5) vemos que el docente C, por un lado, expresa acuerdo para los ítems 1 y 6, lo que indicaría considerar la evaluación como parte del proceso de aprendizaje y como formativa y continua. Por otro lado, detectamos incoherencias en sus respuestas a los ítems 2, 3 y 4, expresando desacuerdo con que el mejor método para evaluar es la prueba escrita y acuerdo a que utiliza la prueba escrita como el único método de evaluación y, al mismo tiempo, con que complementa la prueba escrita con otros métodos. Además, al considerar la forma en que define la evaluación, aparece la representación de la evaluación de producto, es decir, al final del proceso de enseñanza, no como parte del mismo, ya que le asigna el objetivo de determinar en qué medida se han aprendido los contenidos enseñados.

Así, las respuestas dadas por este docente muestran ciertas contradicciones y se vincularían más con un estilo de enseñanza tradicional, centrado en la enseñanza.

En la figura 6 se presentan los resultados del docente C en relación con aprendizaje.

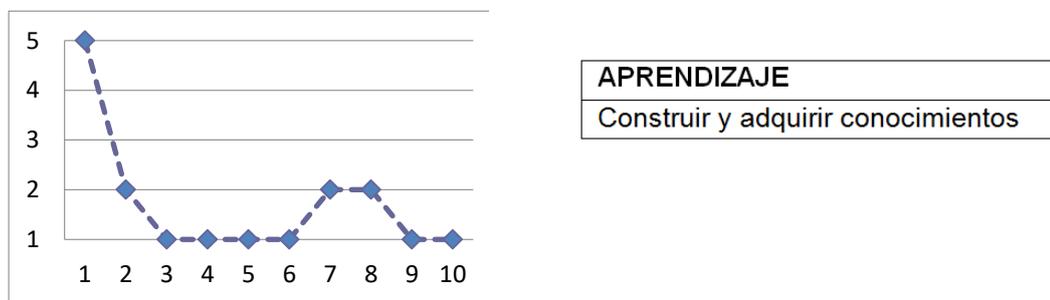


Figura 6. Perfil actitudinal y definición de aprendizaje del docente C  
Fuente: Elaboración propia.

Al analizar el perfil de aprendizaje del docente C vemos que el único ítem para el que no presta su acuerdo es el ítem 1. Con el resto de los ítems está de acuerdo o muy de acuerdo. Estos acuerdos resultan llamativos ya que estas respuestas resultan contradictorios, en tanto algunos ítems (2 y 7) apuntan a un alumno pasivo y que acumule conocimientos y otros (3, 4, 5, 6, 8 y 9) reflejan una concepción de aprendizaje centrada en el alumno y en la construcción de significados. Esta contradicción también aparecería en la misma opinión que da el docente cuando dice que aprender es «construir y adquirir conocimientos», si pensamos la construcción como un proceso activo por parte del alumno y la adquisición como un hecho final pasivo.

Para finalizar con la exposición de los resultados presentamos las respuestas del docente D.

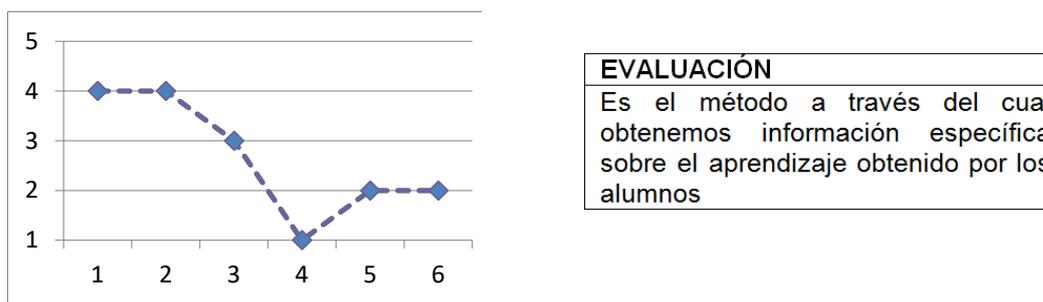


Figura 7. Perfil actitudinal y definición de evaluación del docente D

Fuente: Elaboración propia.

Al observar la figura 7 vemos que el docente manifiesta acuerdo con los ítems 4, 5 y 6 y desacuerdo con los ítems 1 y 2. Esto muestra una idea confusa de evaluación ya que algunas respuestas se contradicen entre sí, por ejemplo, el desacuerdo con el ítem 1 y el acuerdo con el ítem 6, expresa que no evalúa para conocer el proceso de aprendizaje de sus alumnos e introducir las mejoras necesarias pero si utiliza procedimientos de evaluación formativa/continua. Respecto a la definición de evaluación se centra en considerarla como un método que permite obtener información del aprendizaje “obtenido” por los alumnos. De esta manera la representación que puede inferirse muestra a la evaluación como evaluación de producto y como un método que se utiliza sólo para valorar los conocimientos que han adquirido los alumnos. En este sentido, esta representación -similar a la representación del docente C -, se vincularía con un estilo de enseñanza tradicional, centrado en la enseñanza.

A continuación presentamos los resultados obtenidos de las respuestas del docente D para aprendizaje.

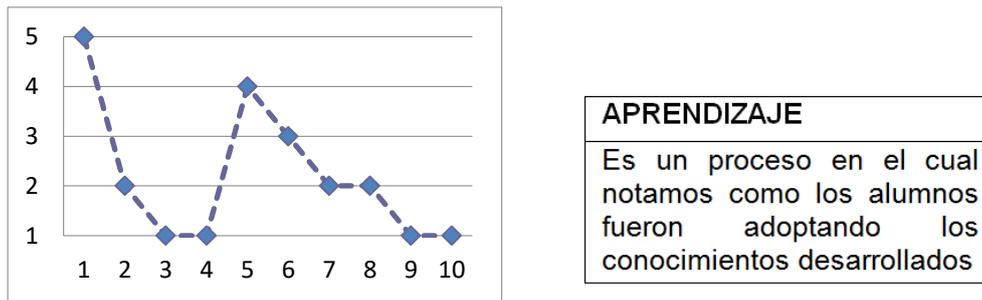


Figura 8. Perfil actitudinal y definición de aprendizaje del docente D  
Fuente: Elaboración propia.

El perfil del docente D (figura 8), muestra que está muy de acuerdo con los ítems 3, 4, 9 y 10 y de acuerdo con los ítems 2, 7 y 8. Estos acuerdos tienden a mostrar que el docente considera la importancia de los saberes de sus estudiantes, que los conocimientos se construyen con la ayuda del profesor y, al mismo tiempo, que considera que el papel del alumno es estar atentos en clase y que aprender es incrementar conocimientos. Respecto a la definición del docente sobre aprendizaje podemos ver que entra en contradicción con algunos de sus acuerdos con los ítems de la encuesta relacionados con la construcción del aprendizaje ya que habla de un proceso que permite mostrar como los alumnos fueron «adoptando» los conocimientos, lo que apuntaría más a acumulación de conocimiento que a la construcción y significación de los mismos por parte de los estudiantes. Así, las respuestas dadas por el docente, aunque muestran ciertas contradicciones, se vincularían con un estilo de enseñanza tradicional, centrado en la enseñanza.

A continuación analizaremos los perfiles actitudinales comparados de los cuatro docentes, referidos a la evaluación. Para el análisis de cada gráfico consideramos los ítems para los que los docentes expresaron acuerdo y algunas relaciones entre las opciones elegidas para los diversos ítems, teniendo en cuenta la presencia de contradicciones entre los acuerdos, indiferencia o desacuerdos.

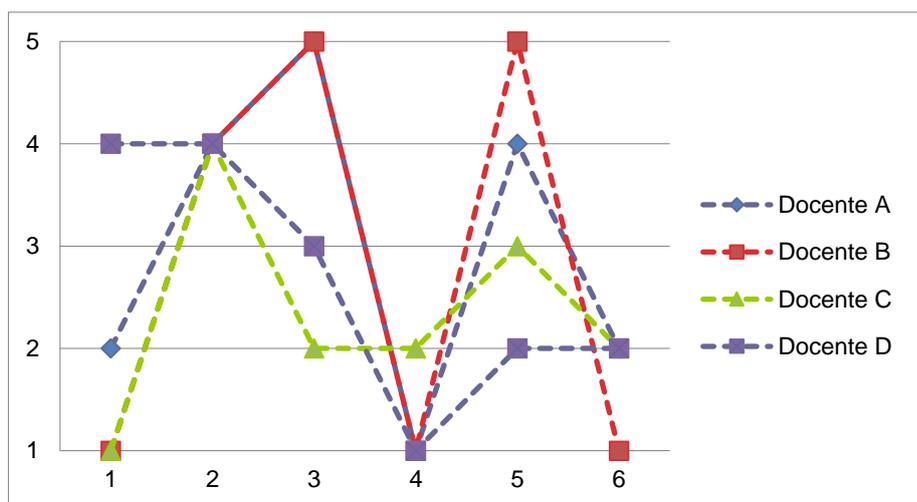


Figura 9. Perfiles actitudinales referidos a evaluación  
Fuente: Elaboración propia.

En la figura 9 podemos observar:

- Para los Ítem 4 y 6 todos los docentes están de acuerdo con complementar la prueba escrita con diferentes métodos de evaluación y realizar una devolución de los resultados tendientes a la mejora de los aprendizajes. Esto mostraría una concepción de evaluación como proceso y como parte del aprendizaje.
- Para el ítem 2 (el mejor método para evaluar a los alumnos es la prueba escrita), todos los docentes expresan desacuerdo, lo cual sería coherente con los acuerdos a los ítems 4 y 6.
- Por otra parte para el ítem 3 (utilizar como método de evaluación únicamente las pruebas escritas) manifiesta indiferencia el docente C y estar de acuerdo el docente D, constituyendo una diferencia de opinión respecto de los ítems 4 y 6, lo que mostraría una contradicción entre las respuestas dadas. En tanto los docentes A y B expresan su total desacuerdo con dicho ítem.
- Para el ítem 1 los docentes B y C están muy de acuerdo, el docente A de acuerdo y el docente D en desacuerdo. Así, la mayoría de los docentes valorarían la evaluación como proceso. Por otro lado, las respuestas al ítem 5 que se refiere a la evaluación sólo como medición de resultados, muestran una dispersión en las opciones elegidas. Los docentes A y B expresan desacuerdo lo que resulta coherente con su respuesta para el ítem 1. En tanto el docente C manifiesta indecisión y el docente D su acuerdo lo que mostraría, para estos profesores, una valoración de la evaluación como producto.

A partir de este análisis vemos que para los docentes A y B sus respuestas no muestran contradicciones entre sí y se relacionan con el estilo de enseñanza centrado en el aprendizaje. En cambio, en el caso de los docentes C y D sus respuestas muestran ciertas contradicciones.

Presentamos la figura 10 con los perfiles actitudinales de aprendizaje.

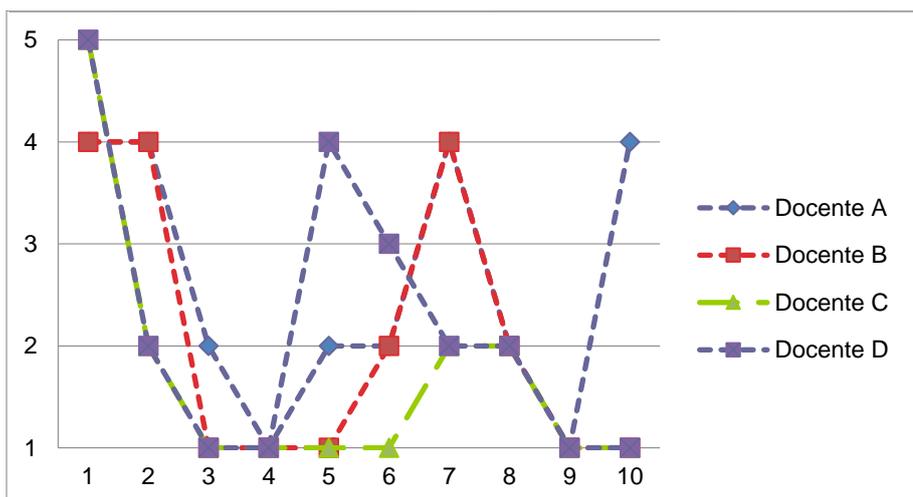


Figura 10. Perfiles actitudinales referidos al aprendizaje

Fuente: Elaboración propia.

Como podemos observar en la figura 10:

- Para los ítem 3, 4, 8 y 9 todos los docentes prestan su acuerdo. Esto muestra una consideración del aprendizaje como construcción a partir de los saberes previos y el contacto con la realidad.
- Respecto al ítem 1 encontramos que todos los docentes expresan su desacuerdo, de esta manera consideran que los aprendizajes no deberían limitarse a los contenidos fundamentales de la disciplina.
- En relación con ítem 5, que se refiere a que el aprendizaje es la construcción personal de significados, el docente D es el único que expresa desacuerdo.
- Respecto al ítem 6 el docente D expresa indecisión, mientras los demás acuerdan con que los conocimientos que adquieren los estudiantes les permiten interpretar la realidad.
- En relación con el ítem 2 los docentes A y B expresan su desacuerdo a considerar el aprendizaje como acumulación de conocimiento, en tanto los docentes C y D manifiestan acuerdo.
- La opinión de los docentes para el ítem 7 está dividida, mientras los docentes A y B están en desacuerdo con que el papel básico de los alumnos en las clases es estar atentos y tomar apuntes, los docentes C y D expresan estar de acuerdo.

Al igual que cuando analizamos los perfiles respecto de evaluación, a partir de este análisis podemos afirmar que las respuestas dadas por los docentes A y B no muestran contradicciones y se relacionan con el estilo centrado en el aprendizaje. En cambio, las respuestas de los docentes C y D muestran contradicciones ya que sus opiniones para algunos ítems se relacionan con un estilo centrado en el aprendizaje, mientras que para otros no.

## **5. Conclusiones**

Los resultados obtenidos nos han permitido diferenciar dos RS de evaluación vinculadas con las posturas teóricas que hablan de la evaluación como medición y como comprensión (Isaza, Mesa y Parra Mosquera, 2006; Santos Guerra, 2005). Así, hemos detectado que la representación de los docentes con un estilo de enseñanza centrado en la enseñanza se acerca a la postura que considera la evaluación como medición, siendo el profesor quien establece los criterios, los aplica e interpreta los resultados para certificar el aprendizaje; mientras que la RS de los docentes que poseen un estilo de enseñanza centrado en el aprendizaje considera a la evaluación como comprensión, como un proceso y no como un momento final y realizándola a través de diferentes instrumentos y medios.

Así mismo, en relación con el aprendizaje, identificamos que los docentes que muestran un estilo centrado en el aprendizaje poseen una representación cuyos elementos son coherentes entre sí y se refieren a un aprendizaje constructivo, considerando la importancia de los saberes previos de los alumnos, sus intereses, su participación activa en el proceso de aprendizaje. En tanto que para los docentes que poseen un estilo centrado

en la enseñanza, su representación de aprendizaje incluye elementos que resultan contradictorios entre sí, al considerar, por ejemplo, la importancia de las ideas espontáneas como punto de partida para el aprendizaje y, al mismo tiempo, estar de acuerdo con algunos ítems del instrumento que apuntan a un alumno pasivo y que acumula conocimientos – estos resultados están en concordancia con otras investigaciones llevadas a cabo (Guirado, 2014; Laudadio, Mazzitelli y Guirado, 2015)-.

Algunos especialistas (Hernández Abenza, 2010; Hidalgo y Hernández-Castilla, 2015; McDonald, Boud, Francis y Gonczi, 2000; Tovar Gálvez, 2008) coinciden en afirmar que han surgido nuevos desafíos en el campo de la evaluación, los cuales requieren respuestas innovadoras, de allí la necesidad de repensar la relación entre evaluación y aprendizaje. En este sentido una evaluación que contribuya con el aprendizaje como proceso de construcción supone no sólo encontrar un método apropiado y usarlo adecuadamente en relación con el contenido con el que se trabaja, sino lograr que el docente reflexione sobre su práctica de manera integral y sobre los supuestos que la fundamentan y las implicancias en el aprendizaje de sus alumnos.

La investigación realizada nos ha permitido no sólo caracterizar cada una de las RS -sobre aprendizaje y evaluación-, sino que hemos estudiado de manera conjunta estas representaciones, pudiendo observar las relaciones entre sus elementos y las vinculaciones con los estilos de enseñanza. Es decir, identificamos las RS de aprendizaje y de evaluación que subyacen a los estilos de enseñanza que guían la práctica docente en el aula. Continuamos trabajando con el desarrollo de estudios que permitan profundizar el análisis y con el desarrollo de actividades que favorezcan los procesos de reflexión con docentes en funciones y estudiantes de profesorado de Ciencias Naturales.

## Referencias

- Alcaraz Salarirche, N. (2015). Aproximación Histórica a la Evaluación Educativa: De la Generación de la Medición a la Generación Ecléctica. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 11-25.
- Abriç, J. C. (2001). *Prácticas sociales y representaciones*. Ciudad de México: Ed. Coyoacán.
- Barrea Andaur, S. (2011). La reflexión docente como dinamizadora del cambio de prácticas en aula. Una experiencia de perfeccionamiento académico en la Universidad Católica Silva Henríquez (UCSH). *Perspectiva Educativa*, 50(1), 31-60.
- Castillo, R., Hernández, J., Munevar, O. y Portilla, M. (2014). Implicaciones de la evaluación continua, a través de rúbricas sobre las prácticas pedagógicas: evidencia empírica y aplicación de análisis multidimensional. *Revista Horizontes Pedagógicos*, 16, 66-77.
- Castorina, A., Barreriro, A. y Toscano, A. (2007). Dos versiones del sentido común: las teorías implícitas y las representaciones sociales. En A. Castorina (Comp.), *Construcción conceptual y representaciones sociales. El conocimiento de la sociedad* (pp. 205-238). Buenos Aires: Miño y Dávila.
- Cea D'Ancona, M. A. (1999). *Metodología cuantitativa. Estrategias y técnicas de investigación social*. Madrid: Editorial Síntesis.
- Celman, S. (2005). ¿Es posible mejorar la evaluación y transformarla en una herramienta de conocimiento? En A. Camilloni et al. (Comp.), *La evaluación de los aprendizajes en el debate didáctico contemporáneo* (pp. 35-36). Buenos Aires: Paidós.

- García Cabrero, B., Loredó Enríquez, J. y Carranza Peña, G. (2008) Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión. *Revista Electrónica de Investigación Educativa*, Especial. Recuperado de <http://redie.uabc.mx/NumEsp1/contenido-garcialoredocarranza.html>
- Gargallo López, B., Suárez Rodríguez, J., Garfella Esteban, P. y Fernández March, A. (2011). El cuestionario CEMEDEPU. Un instrumento para la evaluación de la metodología docente y evaluativa de los profesores universitarios. *Estudios sobre Educación*, 21, 9-40.
- Guirado, A. (2014). *Los modelos didácticos de docentes de Ciencias Naturales de nivel secundario: reconstrucción a partir de sus concepciones y sus prácticas áulicas*. Tesis Doctoral. Universidad Nacional de Cuyo, Mendoza, Argentina.
- Guirado, A., Mazzitelli, C. y Olivera, A. (2013). Representaciones sociales y práctica docente: una experiencia con profesores de Física y de Química. *Revista Orientación Educativa*, 27(51), 1-20.
- Guirado, A., Mazzitelli, C., Olivera, A. y Quiroga, D. (2013). Relaciones entre las representaciones de los alumnos acerca de la enseñanza y el aprendizaje de la Física y de la Química con la práctica docente. *Revista Electrónica de Enseñanza de las Ciencias*, 12(2), 347-361.
- Hernández Abenza, L. (2010). Evaluar para Aprender: hacia una dimensión comunicativa, formativa y motivadora de la evaluación. *Enseñanza de las ciencias*, 28(2), 285-293.
- Hidalgo, N. y Hernández-Castilla, R. (2015). ¿Qué es una evaluación justa? Concepciones de los docentes sobre la de evaluación del aprendizaje de los estudiantes. En AIDIPE (Ed.). *Actas del Congreso Investigar con y para la sociedad*, Vol. 2 (pp.1257-1266). Cádiz: Bubok.
- Kaplan, C. (1997). *La inteligencia escolarizada*. Buenos Aires: Miño y Dávila.
- Kaplan, C. (2003). *Buenos y malos alumnos. Descripciónes que predicen*. Buenos Aires: Aique.
- Laudadio, J., Mazzitelli, C. y Guirado, A. (2015). Representaciones de Docentes de Ciencias Naturales: punto de partida para la reflexión de la práctica. *Actualidades Investigativas en Educación*, 15, 1-23.
- López y Mota, A. D., Rodríguez Pineda, D. P. y Bonilla Pedroza, M. X. (2004). ¿Cambian los Cursos de Actualización las Representaciones de la Ciencia y la Práctica Docente? *Revista Mexicana de Investigación Educativa*, 9(22), 699-719.
- Mazzitelli, C., Aguilar, S., Guirado, A. y Olivera, A. (2009). Representaciones sociales de los profesores sobre la docencia: contenido y estructura. *Revista Educación, Lectura y Sociedad*, 6(6), 265-290.
- Mazzitelli, C. y Guirado, A. (2010). *La enseñanza y el aprendizaje de las Ciencias. Estudios de las representaciones sociales de docentes y futuros docentes en Ciencias*. San Juan: Editorial FFHA – UNSJ.
- McDonald, R., Boud, D., Francis, J. y Gonczy, A. (2000). Nuevas perspectivas sobre la evaluación. *Boletincinterfor*, 149(mayo-agosto), 41-72.
- Moscovici, S. (1961). *La psychanalyse, son image et son public*. París: PUF.
- Moscovici, S. (1979). *El Psicoanálisis, su imagen y su público*. Buenos Aires: Huemul.
- Moral Santaella, C. (2001). La investigación teórico/práctica: Estrategias de formación inicial del profesor. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 4(2), 1-13.
- Pacheco Lora, L. C. (2013). La reflexión docente: eje para promover el cambio representacional de concepciones y prácticas en los docentes. *Zona Próxima*, 19, 107-118.

- Salinas Salazar, M. L., Isaza Mesa, L. S. y Parra Mosquera, C. A. (2006) Las representaciones sociales sobre la evaluación de los aprendizajes. *Revista Educación y Pedagogía*, 18(46), 203-221.
- Santos Guerra, M. A. (2005). Evaluar es comprender. De la concepción técnica a la dimensión crítica. *Revista Investigaciones en Educación*, 5(1), 67-85.
- Torres, A. M. (2013). Representaciones sociales sobre el proceso evaluación desde la mirada de docentes de 1° Año Básico en establecimientos municipalizados urbanos de la comuna de Quilpué. *Estudios Pedagógicos*, 39(1), 285-304.
- Tovar Gálvez, J. C. (2008). Fundamentos y líneas de trabajo propuesta de modelo de evaluación multidimensional de los aprendizajes en Ciencias Naturales y su relación con la estructura de la Didáctica de las Ciencias. *Revista Eureka Enseñanza Divulgación Científica*, 5(3), 259-273.
- Umaña Mata, A. C. (2014). Evaluación de modelos de diseño instruccional: una revisión de literatura. *Innovaciones Educativas*, 16(21), 23-30.

## Breve CV de las autoras

### **Claudia Alejandra Mazzitelli**

Profesora de Enseñanza Media y Superior en Física (UNSJ). Doctora en Educación (UNCuyo). Profesora titular de cátedras de formación docente. Investigadora de la Universidad Nacional de San Juan y del CONICET (Investigador Independiente). Directora del Instituto de Investigaciones en Educación en las Ciencias Experimentales (Facultad de Filosofía, Humanidades y Artes-Universidad Nacional de San Juan). ORCID ID: 0000-0002-1199-4843. Email: mazzitel@ffha.unsj.edu.ar

### **Ana María Guirado**

Licenciada en Psicopedagogía (UCC). Doctora en Educación (UNCuyo). Profesora titular de cátedras de formación docente en la Facultad de Filosofía, Humanidades y Artes-Universidad Nacional de San Juan. Investigadora del Instituto de Investigaciones en Educación en las Ciencias Experimentales (Facultad de Filosofía, Humanidades y Artes-Universidad Nacional de San Juan). Sub-Directora del Departamento de Física y de Química (Facultad de Filosofía, Humanidades y Artes-Universidad Nacional de San Juan). ORCID ID: 0000-0002-0974-3981. Email: aguirado@ffha.unsj.edu.ar

### **María Julieta Laudadio**

Licenciada en Ciencias de la Educación (UCA). Doctora en Educación (UNCuyo). Investigadora del Instituto de Investigaciones en Educación en las Ciencias Experimentales (Facultad de Filosofía, Humanidades y Artes-Universidad Nacional de San Juan). Becaria del CONICET. ORCID ID: 0000-0003-4049-6752. Email: julietalaudadio@gmail.com

## Efecto de la Integración de las Artes Visuales en el Desarrollo de la Lectura y la Escritura

### Effect of the Integration of Visual Arts in Reading and Writing Skills

Wanda I. Padilla Martínez \*  
Jaime Ortiz-Vega

Pontificia Universidad Católica de Puerto Rico – Puerto Rico

La educación actual enfrenta grandes desafíos. Anualmente, en Puerto Rico, gran parte de los estudiantes toman pruebas estandarizadas en las que deben demostrar dominio de las competencias en el área académica de español, entre otras. Este artículo presenta una investigación cuyo propósito fue determinar el efecto de la integración de las artes visuales, durante diez semanas, en el desarrollo de las destrezas de lectura y escritura en estudiantes de tercer grado de escuelas públicas del suroeste de Puerto Rico. Se utilizó un enfoque cuantitativo con un diseño cuasi-experimental para determinar la diferencia entre: las puntuaciones de la posprueba de lectura y de escritura, respectivamente, de los estudiantes a los que se le aplicó el tratamiento y a los que no se le aplicó; las puntuaciones de la preprueba y la posprueba de lectura y de escritura, respectivamente, obtenida por los estudiantes de cada grupo. La muestra de 21 estudiantes de tercer grado se seleccionó por conveniencia, en dos escuelas similares. De éstos, 14 formaron parte del grupo control (GC) y 17 del grupo experimental (GE). Los datos se recolectaron mediante la pre y posprueba de lectura y la pre y posprueba de escritura creada por la investigadora. La información recopilada se analizó mediante estadística inferencial paramétrica (t test, ANOVA, regresión lineal) y estadística no paramétrica (Mann Whitney, Wilcoxon). Los resultados evidenciaron diferencias significativas entre GC y GE, en los promedios de las puntuaciones de las prepruebas y de las pospruebas de lectura y escritura, respectivamente. Estos hallazgos sugieren que la integración de las artes visuales tuvo un efecto significativo en el desarrollo de las destrezas de lectura y escritura.

**Palabras clave:** Pruebas estandarizadas; Destrezas del lenguaje; Desarrollo de lectura; Destrezas de escritura; Educación en arte; Estrategias de aprendizaje.

Today's education faces major challenges. Annually, in Puerto Rico, most of the students take standardized tests in which they must demonstrate mastery in the competences in the academic area of Spanish, among others. This article presents an investigation whose purpose was determine the effect of the integration of visual arts, for ten weeks, in the development of Reading and writing skills in third grade students in public schools in southwest of Puerto Rico. A quantitative approach was used with a quasi-experimental design to determine the difference between: the post-test scores of reading and writing, respectively, of the students to whom the treatment was applied and to whom it was not applied; the pre-test and post-test scores of reading and writing, respectively, obtained by the students of each group. The sample of 21 third graders was selected for convenience in two similar schools. Of these, 14 were part of the control group (CG) and 17 of the experimental group (EG). The data were collected through the pre and post reading test and the pre and post writing test created by the researcher. The information collected was analyzed by parametric (t test, ANOVA, lineal regression) and non-parametric (Mann Whitney, Wilcoxon) inferential statistics. The results evidenced significant differences between CG y EG, in the averages of the pre and post Reading and Writing test scores, respectively. These findings suggest that the integration of the visual arts had a significant effect on the development of Reading and writing skills.

**Keywords:** Standardized Tests; Language Skills; Reading Skills; Writing Development; Art Education; Learning Strategies.

---

\*Contacto: wandapadilla@pucpr.edu

issn: 1989-0397

www.rinace.net/riee/

https://revistas.uam.es/riee

Recibido: 17 de agosto de 2017

1ª Evaluación: 11 de octubre de 2017

Aceptado: 8 de diciembre de 2017

## 1. Introducción

En una era en que se ha logrado, a pasos agigantados, un avance en los diversos campos del saber se buscan maneras nuevas de fortalecer el aprendizaje de modo que éste trascienda a las disciplinas (Krishna, 2004; Santiago-Villafañe y Figarella-García, 2012). En este contexto, el Departamento de Educación de Puerto Rico (DEPR) se enfrenta al reto de transformar las escuelas con el propósito de mejorar el aprovechamiento académico de los estudiantes (DEPR, 2014a). Según lo establece el DEPR, la enseñanza de las artes del lenguaje en español es una de las áreas de prioridad. De acuerdo con DEPR (2015a) siendo el español la lengua materna, es el principal medio de comunicación que regula la interacción entre los seres humanos que hablan esta lengua y, por consiguiente, es el lenguaje con que los estudiantes aprenden los conocimientos de las demás áreas del saber. El dominio del idioma, según el DEPR (2014b), provee las habilidades para el devenir de las nuevas situaciones en un mundo dinámico al servir como apoyo una formación integral que tiene como resultado, educandos activos, críticos y participativos.

En Puerto Rico, los estudiantes de escuela pública pasan por un proceso de medición anual denominado Pruebas Puertorriqueñas de Aprovechamiento Académico (PPAA). A partir del año 2016, estas pruebas se denominan Medición y Evaluación para la Transformación Académica de Puerto Rico (META-PR). Estas pruebas, alineadas a los estándares académicos establecidos por el DEPR, están diseñadas para evaluar el aprovechamiento académico en diferentes materias, entre estas, español. La ley *No Child Left Behind* (NCLB, 2002) requiere que los estudiantes matriculados en los grados 3er, 4to, 5to, 6to, 7mo, 8vo y 11vo, participen de este programa de medición. El cuadro 1 presenta información relacionada con el desempeño de los estudiantes del DEPR en las PPAA durante los años 2012 a 2015.

Se observa en el cuadro 1, que los estudiantes en los años 2012 a 2015 no alcanzaron la meta anual estipulada por las leyes federales que rigen la educación en Puerto Rico (PR). De acuerdo con el DEPR (2015b) los estudiantes han tenido dificultad en demostrar dominio en el área de español. La situación se complica, ya que en el área académica de español las metas anuales son más elevadas cada año (DEPR, 2014a). En el cuadro 2 se presentan los objetivos medibles anuales (*Annual Measurable Objectives*, o AMO) que establece el sistema de educación diferenciada en el Plan de Flexibilidad (DEPR, 2014a; USDE, 2015) para la materia de español, para los grados de tercero a octavo.

Cuadro 1. Desempeño de los estudiantes en PPAA (DEPR, 2015b)

	ESPAÑOL %	PROGRESO %	INGLÉS %	PROGRESO %	MATEMÁTICAS %	PROGRESO %	TOTAL %	PROGRESO %
<b>2012-2013</b>	47,88		42,61		28,78		39,75	
<b>2013-2014</b>	45,71	-2,17	41,35	-1,26	30,03	1,24	39,02	-0,73
<b>2014-2015</b>	43,55	-2,16	41,91	0,56	29,64	-0,39	38,36	-0,66

Fuente: Elaboración propia.

Cuadro 2. Objetivos medibles anuales (AMO) para español, 3er a 8vo grado (DEPR, 2014a)

	LÍNEA BASE						
	2011- 2012	2012- 2013	2013- 2014	2014- 2015	2015- 2016	2016- 2017	2017- 2018
	%	%	%	%	%	%	%
Todos los estudiantes	47,7	52,1	56,4	60,8	65,1	69,5	73,9

Fuente: Elaboración propia.

El panorama actual de la educación pública, para los grados K-12 en Puerto Rico, es el resultado de no haber podido cumplir con la ley federal *Elementary and Secondary Education Act* (ESEA)(1965) según enmendada por la ley 107-110 (2002) conocida como *No Child Left Behind* (NCLB). La mencionada ley requería que los estudiantes lograran el 100 por ciento de proficiencia académica para el año 2014, sin embargo, el objetivo no se logró. No obstante, en el 2011, el Departamento de Educación de Estados Unidos (DEEU) ofreció a los estados y territorios la oportunidad de hacer un plan para acogerse a las dispensas otorgadas a requisitos particulares de la ley ESEA. El DEPR presentó un Plan de Flexibilidad al DEEU con el objetivo de lograr el dominio de los estándares y la oportunidad de mejorar el desempeño de los estudiantes, según sus necesidades diferenciadas (DEPR, 2014a). Asimismo, indica que la meta de ese sistema de rendición de cuentas es aumentar la transparencia en los procesos educativos y lograr que los grupos de interés se involucren en el proceso de transformar las escuelas, de manera que esto resulte en un sistema de escuelas públicas que satisfagan las necesidades de los estudiantes.

De acuerdo con el Informe del Perfil Escolar (DEPR, 2015b) los estudiantes de tercer grado de las escuelas públicas de PR en las cuales se administró las PPAA en el año escolar 2014-2015, el 57,4 % obtuvieron puntuaciones equivalentes a avanzados y proficientes en la parte de español. En el cuadro 3 se presenta el nivel de proficiencia, en el área de español, de los estudiantes de tercer grado comparado con los años anteriores y con los AMO establecidos para cada año.

Cuadro 3. Proficiencia en español en las PPAA, según DEPR (2015b)

AÑOS ESCOLARES	3ER GRADO %	PROGRESO %	AMO 3ER-8VO GRADO %	CUMPLIÓ META
2012-2013	55,6	-----	52,1	NO
2013-2014	55,1	-0,5	56,4	NO
2014-2015	57,4	2,3	60,8	NO

Fuente: Elaboración propia.

Si se toma en consideración que la meta anual, según la ley 107-110 (2002), para esta materia era de 83,18% y que, en el nuevo sistema diferenciado, los objetivos anuales medibles establecidos, era de 60,8%, según el cuadro 3, se observa que los estudiantes no dominaron las destrezas del lenguaje, tales como lectura, escritura y comprensión lectora, en los porcentajes esperados.

En este contexto, cabe preguntarse, ¿cuáles acciones se deben tomar para revertir ese estado de situación?, ¿cuál es la probabilidad de que mediante la integración de otras áreas curriculares en la enseñanza de las artes del lenguaje en español se pueda lograr resultados

conducentes a cumplir con las metas establecidas? Si fuera así, ¿cuál es el efecto de la integración de las artes visuales en otras áreas curriculares, en particular el español? De acuerdo con Fiske (1999) de varias investigaciones se desprende que mediante la educación artística se puede transformar el proceso educativo al provocar, entre otros, el aprendizaje por descubrimiento, el mejoramiento de las relaciones interpersonales y la iniciativa de aprender. Según expone la política pública de la organización y la oferta curricular del programa de Bellas Artes en los niveles elemental y secundario de las escuelas públicas (DEPR, 2013), ésta provee al estudiante experiencias de comunicación universal que son necesarias para desenvolverse en el mundo del siglo XXI. Además, Maeso-Rubio (2003) sostiene, refiriéndose a la educación artística que, entre las diversas áreas de enseñanza, es una de las que posee unas características especiales que la diferencian de otras disciplinas.

De acuerdo con lo antes expuesto, surgen otras interrogantes tales como: ¿Cuál es el desempeño de los estudiantes en las destrezas de lectura y escritura? ¿Cuál es el rol de la lectura y la escritura en el desempeño académico? ¿Cuál es la importancia de las artes visuales en la educación?

Si se examinan los orígenes de la educación en bellas artes, según Eisner (1972) se notará que éstas han sido catalogadas como un área de estudio que no va más allá de lo que es el producto estético. Asimismo, en los últimos tiempos, las culturas han considerado las artes como una materia de poca importancia en una sociedad llena de riqueza que, además, posee mucho tiempo disponible (Sousa, 2002).

En Puerto Rico, no fue hasta el 1955, que se creó el Programa de Artes Visuales (DEPR, 2003a). Desde ese momento, la educación artística ha sido parte del currículo en las escuelas públicas de la Isla. Este programa pasó por procesos de consolidación y separación hasta que, en el 2000 su estructura se fortaleció a través de estándares académicos en los diferentes niveles de enseñanza. El DEPR (2003b) documenta que es a partir de esta fecha en que los enfoques de enseñanza se actualizaron y se alinearon a las nuevas tendencias educativas que favorecerían la implantación de estrategias para la expresión artística.

## **2. Fundamentación teórica**

De las investigaciones se desprende que el proceso natural de alfabetización se inicia antes de la instrucción formal en la escuela (Goodman y Altwerger, 1981; Harste, Burke y Woodward, 1982; Mass, 1982; Ruiz, 1998). Es, precisamente esta conceptualización temprana del lenguaje impreso como proceso de comunicación, un elemento clave para el desarrollo eficaz de las destrezas de lectura y la escritura desde las etapas iniciales en el aprendizaje formal hasta la realización del lector hábil para toda la vida (Kontos y Huba, 1983). En este sentido, Fons (2010) afirma que es necesario propiciar un ambiente rico en experiencias de lectura y escritura que permita que los niños entren en contacto con una variedad de situaciones que se resuelvan mediante actos en los que tendrán que leer y escribir.

Goodman (2003) señala que el lenguaje se aprende en el proceso de su uso, mediante la participación activa y repetida en actividades que impliquen hablar y escribir. De esta manera, según el autor, se llevará a cabo un proceso natural del aprendizaje del lenguaje,

denominado *lenguaje integral*; lo cual significa que el aprendiz va apropiándose del lenguaje desde su nacimiento y a través de todas las experiencias que impliquen el desarrollo de la actividad comunicativa.

Leer, según Solé (2010) es un proceso dialéctico entre un texto y un lector, proceso en el que éste aporta su disposición emocional y afectiva, sus propósitos, su experiencia, su conocimiento del mundo y del tema. Es esa experiencia, de acuerdo con el autor, la que permite comprender y construir un significado respecto a lo leído. Esta comprensión, será más profunda en la medida en que el estudiante pueda asociar su experiencia como lector con los conceptos que surgen del mundo de lo cotidiano (Freire, 2004).

Según Schneuwly (2010) el lenguaje escrito, cuya construcción es una de las tareas prioritarias del DEPR, supone la reorganización de distintas capacidades que forman una nueva esencia interior. También afirma el autor, que como parte de este proceso, que hace visible el idioma en uso, se requiere la utilización de diversas funciones intelectuales desde las más elementales a las más complejas para crear una combinación específica a través de las que se construye una nueva unidad. Tal es así, que Vygotsky (1995) sostiene que la escritura es el álgebra del lenguaje.

La práctica de la lectura y la escritura son procesos que se nutren recíprocamente (Rolón, 2013). No obstante, cuando se separan con la intención de fortalecer uno más que otro, según señala Dollins (2001) se debilita esa relación. De acuerdo con este autor, que llevó a cabo una investigación cualitativa sobre la escritura, en la que se ofreció ayuda remediadora a estudiantes rezagados en lectura. La evidencia reveló una transferencia del contenido utilizado en el proyecto cuando la conexión entre lectura y escritura era explícita.

En una investigación similar, Coady (2007) demostró la relación existente entre la lectura y la escritura. En medio de un panorama en el que los fondos que recibían las escuelas daban un particular énfasis a la lectura mediante el proyecto *Leer primero*, el autor decidió retar al sistema para lo cual llevó a cabo la investigación *Que ningún escritor se quede atrás*. Utilizó como marco de referencia, las teorías constructivistas de Vygotsky (1995) y Dewey (1934). Además, el enfoque de Graves (2004) que postuló que “la escritura es la realización de la lectura” (p. 4). Graves reclutó a un maestro voluntario para enseñar a un grupo dando énfasis a la enseñanza de la escritura. Como resultado de la investigación, se identificó la lectura y escritura como procesos que están conectados en la alfabetización.

En otra investigación, Reyes-Tosta (2013) expuso que “el mismo gobierno puertorriqueño ha admitido que su sistema de educación ha fracasado en su intento de lograr que los estudiantes dominen las destrezas de comunicación verbal y escrita del español” (p. 44). Según el autor, los resultados de las PPAA, que ofrece el DEPR, muestran que los estudiantes carecen del dominio de las destrezas básicas del lenguaje. El autor recomienda que, para lograr el éxito de éstos en ambientes educativos avanzados, se debe fomentar la lectura y escritura, y a su vez, la confianza y la competencia necesaria para desarrollar las mismas.

### **2.1. Las artes visuales y el lenguaje**

Deasy (2002) presentó un compendio, el cual contiene 62 investigaciones sobre las relaciones determinantes entre el aprendizaje en las bellas artes, el aprovechamiento académico y el desempeño social. Mediante un análisis de los resultados obtenidos en cada

una de estas investigaciones, Catterall (2002) sostiene que se comprobó que aquellas escuelas que incorporaron la enseñanza y aprendizaje alrededor de las artes visuales en el currículo experimentaron amplios logros tanto en el aprovechamiento académico como en la actitud positiva que se generó en los estudiantes respecto a la motivación para aprender. De acuerdo con el autor, de estas investigaciones se infiere que cada disciplina de las artes visuales se asocia con logros importantes. Establece, por ejemplo, que el dibujo ayudó a desarrollar las destrezas de escritura, mientras que el entrenamiento visual ayudó a la interpretación de textos. Del mismo modo, el estímulo de la imaginación es fundamental en el desarrollo intelectual ya que se trata de ver las cosas de una forma diferente a como son en realidad, siendo esto “lo que hacen los científicos y los artistas; perciben lo que es, pero imaginan lo que podría ser y luego usan su conocimiento, sus aptitudes técnicas y su sensibilidad en lo que han imaginado” (Eisner, 2004, p. 243).

Investigadores tales como Madura (1998), Fiske (1999), Deasy (2002), Catterall (2002), Eisner (2004), Andrzejczak, Trainin y Poldberg (2005), Sanders (2006), Senitt (2006), Kay (2008) y Rachlin (2008) sugieren los múltiples beneficios que las artes aportan al currículo de las escuelas. Andrzejczak, Trainin y Poldberg (2005) llevaron a cabo una investigación basada en una estrategia de enseñanza en la que utilizaron las artes visuales como parte de un proceso de pre-escritura. El mismo se desarrolló como parte de un proyecto de investigación acerca del efecto de las estrategias de las artes en el aprovechamiento académico de los estudiantes. Los autores citados, concluyeron que la incorporación de las artes visuales en el proceso de escritura trajo consigo el enriquecimiento en el lenguaje metafórico y sensorial del estudiante y, a su vez, la escritura añadió al trabajo artístico una nueva dimensión de la idea original.

Sanders (2006) llevó a cabo una investigación de enfoque fenomenológico, en la que participaron seis estudiantes de cuarto grado que estuvieron inmersos en un currículo de arte y escritura. A través de esta investigación, la autora exploró la manera en que se relacionan la composición artística con la composición escrita, y viceversa. Además, mediante el análisis de la experiencia del estudiante midió la efectividad de la enseñanza a través de modelos múltiples. Para esto, utilizó también la lectura mediante la clasificación por unidades temáticas, tales como poesía, historia, artistas, técnicas, entre otros.

En otra investigación, Rachlin (2008) identificó la necesidad de reforzar la destreza de escritura, ya que los estudiantes de escuela intermedia obtenían puntuaciones bajas en las pruebas estandarizadas y en la escuela superior sucedía lo mismo con las pruebas de cuarto año; en ambas, el problema era la presencia de errores al escribir. El investigador enfocó el diseño de un currículo, que denominó “Tengo algo que decir”, en aquellos estudiantes cuya primera lengua no era el inglés. Éste se basó en el estudio de murales mexicanos para enseñar la escritura persuasiva. Se utilizó el arte como un organizador visual para la elaboración de sus ensayos. Las conclusiones fueron que, así como el arte es personal, la escritura también lo es. El autor, consideró que es mediante un enfoque personal que se desarrollan exitosamente ambas.

## ***2.2. El lenguaje, las artes visuales y el aprendizaje significativo***

La evolución de los medios masivos de comunicación, en particular *Internet*, permiten al ser humano acceder a todo tipo de información fácilmente. Ante esto, la escuela del siglo XXI tiene ante sí el reto de desarrollar estudiantes educados y sensibles, capaces de hacer uso y transferir la información a conocimientos que puedan ser utilizados en situaciones

reales. De acuerdo con Álvarez (2006) el ser humano aprende de forma natural como respuesta a las experiencias que ha vivido a lo largo de los años. Es por eso, que no se debe desligar el aprendizaje de las emociones y sentimientos que surgen en el contexto social en que se desarrolló el nuevo conocimiento (Damasio, 2003). De esto se puede sugerir, que cuando la memoria del aprendizaje nuevo está asociada con una experiencia gratificante, el aprendizaje será perdurable (Dewey, 1934; Sousa, 2002).

Una idea común es que un trabajo artístico se identifique como algo aparte de la experiencia humana, no obstante, la experiencia ocurre en cada interacción que hace el ser humano con su ambiente (Dewey, 1934). La experiencia directa, por tanto, propicia conocimientos, siendo ésta una manera de adquirirlos. Según Sousa (2002) es el mecanismo que encauza la enseñanza a incorporar el nivel de significado atribuido al nuevo conocimiento. Entonces, si este aprendizaje se da en condiciones normales, será difícil de olvidar.

Las destrezas del lenguaje se desarrollan en el aprendiz desde el nacimiento. No obstante, éste añade los nuevos conocimientos a los que ya ha adquirido (Ausubel, 2002) para lograr llegar a conclusiones nuevas, aplicándolo, a su vez, a nuevas situaciones. De esta manera el niño que lee y escribe, cada vez más, cobrará confianza y seguridad en su propio aprendizaje.

### ***2.3. Marco conceptual y teórico***

El marco teórico de la investigación incorpora las definiciones y teorías de: Rabkin y Redmond (2004), para el constructo de artes visuales; Goodman (2003), para los constructos de lectura y escritura; y Ausubel (2002) para el constructo de aprendizaje significativo. A continuación, se presenta una descripción de las mismas de acuerdo con la relación de las variables investigadas.

#### ***2.3.1. Artes Visuales***

Rabkin y Redmond (2004a) expusieron su perspectiva acerca de las artes visuales en la educación. La enseñanza y la integración curricular de las artes visuales, según los autores, son dos conceptos muy diferentes. Diferenciaron la integración como una estrategia eficaz para provocar cambios positivos en estudiantes, particularmente, aquéllos pertenecientes a la población por debajo del nivel de pobreza y en desventaja social y académica. Asimismo, los autores destacaron los beneficios cuando se incorporan, de manera estratégica e interdisciplinaria las artes visuales en otras materias. Entre éstos, identificaron la transferencia en el aprendizaje y, respecto a las artes visuales, el uso de la dimensión emocional, social y sensorial de éstas para convertirse en adultos exitosos, el aumento en el aprendizaje y la conexión del contenido con las experiencias personales y necesidades del estudiante para un mejor entendimiento del mundo.

#### ***2.3.2. Lectura y escritura***

De acuerdo con Goodman (2003), el aprendizaje del lenguaje es un proceso mediante el cual se da la invención y la convención. El ser humano, según el autor, inventó el lenguaje para poder comunicarse, para aprender y para pensar; luego, al unirse cada individuo con sus invenciones personales y conformar la sociedad, se crean las convenciones sociales. De esta manera, las invenciones personales y convenciones sociales del lenguaje interactúan entre sí.

El lenguaje escrito, según afirma Goodman (2003), es el resultado del uso del lenguaje oral. No obstante, de acuerdo con el autor, la lectura y la escritura son procesos que se desarrollan a la par ya que se construyen a partir de lo que se conoce; los niños observan lo impreso, responden a éste otorgándole significado, lo imitan y llegan al lenguaje escrito convencional.

### 2.3.3. Aprendizaje significativo

Ausubel (1962), propuso la teoría del aprendizaje significativo definiéndola como el proceso a través del cual una persona incorpora la nueva información de manera que ésta se relacione con la estructura cognitiva previamente existente. A este proceso lo denominó asimilación. Según Ausubel (1962, 2002), la estructura cognitiva previamente existente es el conjunto de conceptos e ideas que se encuentran en el aprendiz, así también como la organización de las mismas.

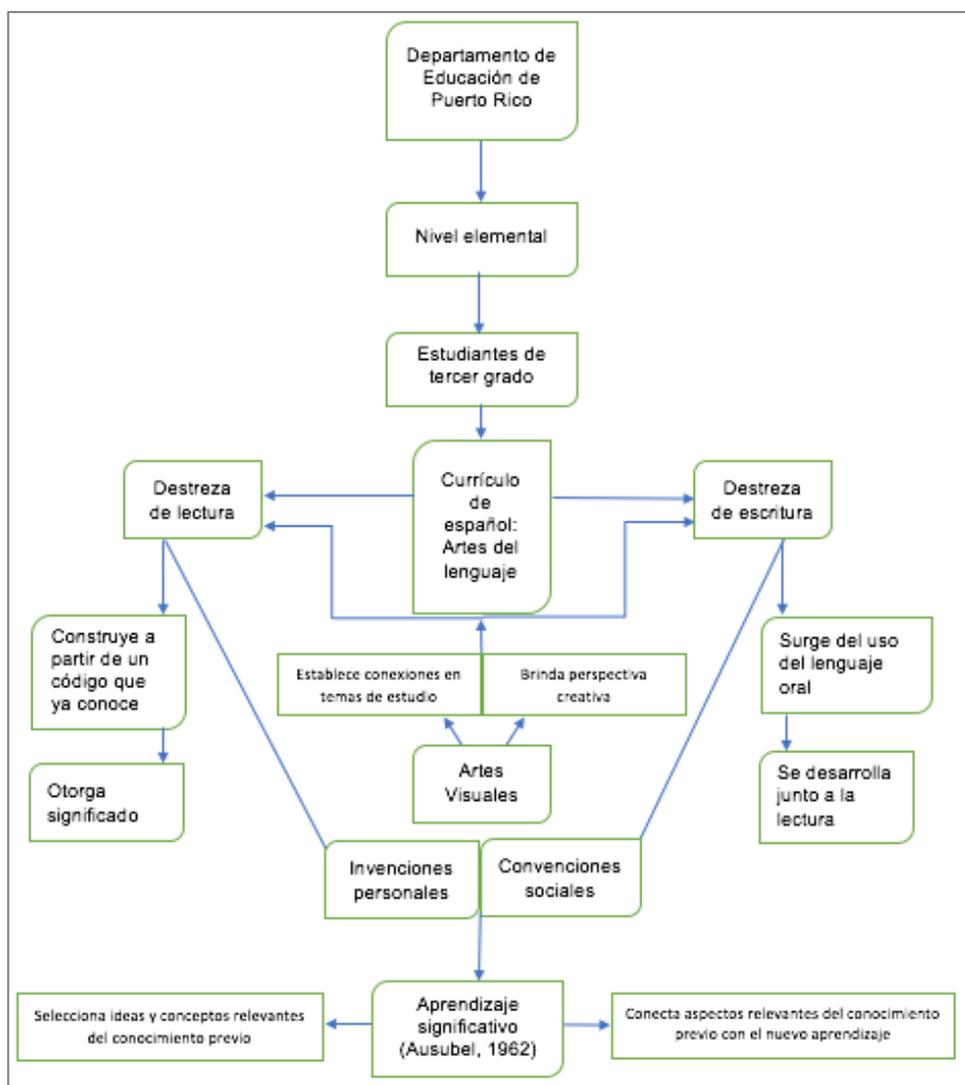


Figura 1. Marco conceptual teórico de la integración de las artes visuales en el currículo de español

Fuente: Elaboración propia.

De acuerdo con el autor, para que se lleve a cabo este proceso: primero, debe tener motivación; segundo, debe poseer conocimientos previos; y, tercero, debe construir nuevos significados. Este proceso ocurre, según el autor, cuando el individuo selecciona las ideas y conceptos más relevantes y los añade al nuevo conocimiento, lo cual hace que la totalidad sea comprensible en cada experiencia de aprendizaje. No obstante, indican Ausubel (1962, 2002) y Ausubel, Novak y Hanesian (2009), que el aprendiz tiene que mostrar la capacidad de reconocer qué aspectos de lo que ya conoce son relevantes en la experiencia de aprendizaje actual.

### 3. Métodos

Para llevar a cabo esta investigación se utilizó un enfoque cuantitativo con un diseño cuasi-experimental. Creswell (2014) indica que el enfoque cuantitativo se debe utilizar, entre otras, en investigaciones para las que se requiera determinar el efecto en las variables de estudio luego de aplicar un tratamiento. Asimismo, Salkind (1999) sostiene que los diseños cuasi-experimentales permiten identificar las diferencias entre las variables, antes y después del tratamiento; a su vez, facilitan a los investigadores explorar preguntas que no podrían contestarse de otro modo. De acuerdo con McMillan y Schumacher (2010), el propósito de un diseño de investigación es proporcionar respuestas apropiadas a las preguntas de investigación planteadas.

Para ilustrar el mismo, el cuadro 4 muestra el diseño de investigación, en el que participó un grupo experimental y un grupo control. Los dos grupos fueron expuestos a cinco lecturas y a ejercicios de escritura durante un período de diez semanas.

Cuadro 4. Diseño de investigación

GRUPOS	PREPRUEBA	TRATAMIENTO	POSPRUEBA
GE (Grupo experimental)	O <sub>1</sub>	X	O <sub>2</sub>
GC (Grupo control)	O <sub>1</sub>		O <sub>2</sub>

O<sub>1</sub> = preprueba, O<sub>2</sub> = posprueba, X = tratamiento.

Fuente: Elaboración propia.

El grupo experimental (GE) leyó las historias y escribió mediante tareas en las que, para propósitos de esta investigación, se integró las artes visuales a través de un segmento de integración curricular diseñado para estos fines. El grupo control (GC) leyó las mismas historias y escribió mediante tareas en las que se llevó a cabo el desarrollo de estas destrezas de la manera tradicional, sin incorporar las artes visuales.

#### 3.1. Muestra

La muestra de esta investigación fue de 21 estudiantes de tercer grado los cuales fueron seleccionados por conveniencia. Los dos grupos fueron heterogéneos, en los que se incluyó los estudiantes de la corriente regular y los estudiantes de educación especial de dos escuelas públicas rurales ubicadas en el suroeste de Puerto Rico.

#### 3.2. Instrumentos de investigación

Los instrumentos de investigación fueron diseñados por la investigadora con la colaboración de una persona con vasta experiencia en la enseñanza y supervisión en el

área académica de español. Para la elaboración de éstas, se consideraron los estándares de contenido y las expectativas para el tercer grado del programa de español del DEPR.

### *3.2.1. Preprueba y posprueba de lectura*

La prueba de comprensión lectora contenía dos lecturas. Se incluyó 18 preguntas de respuestas múltiples, basadas en las lecturas. Cada pregunta constaba de tres alternativas como probable respuesta: una respuesta correcta y dos distractores. Se proveyó una hoja a cada estudiante para que marcara la respuesta correcta al ennegrecer la burbuja en la cual se encontraba la letra seleccionada. Mediante un estudio piloto, esta prueba obtuvo un coeficiente de confiabilidad alfa de Cronbach con valor de 0,82.

### *3.2.2. Preprueba y posprueba de escritura*

La prueba de escritura creativa consistió en la presentación de dos imágenes de obras de arte, las que se solicitó al estudiante observar. A continuación, se pidió al mismo que basándose en lo que observó, escribiera un párrafo de no menos de ocho oraciones. Dicho párrafo se evaluó mediante el uso de la Rúbrica para evaluar la destreza de escritura, la cual contenía 13 criterios para cada una de las dos imágenes presentadas. Mediante un estudio piloto, esta prueba obtuvo un coeficiente de confiabilidad alfa de Cronbach con valor de 0,95.

### *3.2.3. Segmento de integración curricular*

Este segmento de integración curricular se basó en tareas de integración artística que fueron diseñadas para complementar las lecturas incluidas en el currículo de adquisición de la lengua el cual incluye la enseñanza del español como lenguaje materno integrado con la historia de Puerto Rico. Por consiguiente, para la construcción del mismo se tomó en consideración el contenido del Mapa curricular de Adquisición de la Lengua para tercer grado y la comunicación constante con la maestra de este grupo. De esta manera, durante diez semanas se incluyó actividades de los programas de español y de artes visuales apropiadas para su nivel.

## **3.3. Procedimiento**

La investigación fue autorizada por el DEPR y la Junta de Revisión Institucional de la Pontificia Universidad Católica de Puerto Rico. Se explicó la investigación al director, maestros, estudiantes y padres participantes, de las escuelas seleccionadas. Se obtuvo el consentimiento y asentimiento informado de los participantes y tutores de éstos.

### *Administración de preprueba de lectura y de escritura*

Una vez, obtenidos los documentos requeridos para realizar la investigación, se administró las prepruebas de lectura y de escritura, respectivamente, al grupo control (GC) y grupo experimental (GE).

### *Administración de tratamiento a través de la integración de las artes visuales*

El tratamiento constó de la integración de las artes visuales a la enseñanza de la adquisición de la lengua de español por 30 minutos, al GE, durante diez semanas. Para dar dirección a esta etapa, la investigadora diseñó un segmento de integración curricular a través del cual tomó en consideración el Mapa curricular de Adquisición de la Lengua para tercer grado. Específicamente, la primera unidad temática es *Soy Puerto Rico*:

*autografías y narrativas personales* a través de la cual los estudiantes estuvieron expuestos a actividades que los dirigieron a comprender que las biografías incluyen eventos y detalles de la vida de la persona. A medida que, la maestra de español los expuso a la lectura de autobiografías y a la identificación de aspectos tales como: la narrativa y el propósito del autor, la secuencia, los intereses, las capacidades y los valores personales, entre otros, la maestra de artes visuales le enseñó acerca de los autorretratos. Al mismo tiempo, identificaron similitudes entre éstos y las autobiografías. Además, establecieron la diferencia entre retrato y autorretrato de la misma manera que diferenciaron una biografía de una autobiografía. En otra actividad, con la maestra de español prepararon una línea del tiempo del pelotero Roberto Clemente, mientras que con la de artes visuales prepararon una línea del tiempo del pintor Francisco Oller, ambos personajes son considerados puertorriqueños famosos. Mientras leyeron *Un pintor puertorriqueño sobresaliente: su vida y proyecciones* (Cruz y Cruz-López, 1979) también observaron la obra del pintor José Campeche, el cual es el personaje principal de la misma. De igual manera, los estudiantes dibujaron, pintaron e hicieron grabados durante esta etapa de integración.

#### *Administración de posprueba de lectura y de escritura*

Al finalizar las diez semanas, los estudiantes de los dos grupos contestaron las pospruebas de lectura y de escritura, respectivamente, siendo los mismos instrumentos que fueron utilizados para la preprueba.

## **4. Resultados y Conclusiones**

El análisis estadístico incluyó la prueba Mann-Whitney y la prueba t para grupos independientes, para analizar las puntuaciones de la preprueba y las puntuaciones de la posprueba, respectivamente. También se realizó la prueba Wilcoxon y la prueba t para muestras pareadas, para analizar las puntuaciones y determinar las diferencias entre los grupos. Para determinar los efectos del tratamiento en el GE se llevó a cabo el análisis de regresión lineal simple. Las hipótesis se sometieron a un nivel de significancia o alfa ( $\alpha$ ) de ,01 y ,05. Todos los análisis se realizaron a través del uso del programa estadístico SPSS Versión 24. A continuación, se presenta el cuadro 5 en la cual se incluye el promedio de las puntuaciones obtenidas por los estudiantes participantes de los dos grupos, para la posprueba de lectura y de escritura, respectivamente.

El cuadro 5 presenta los datos que se obtuvieron de la posprueba de lectura y de escritura. En la posprueba de lectura, las alternativas para responder a las preguntas, según el rango, son: 3 = totalmente correcta, 2 = parcialmente correcta y 1 = totalmente incorrecta. Esto implica que el promedio más alto para esta prueba es de 3. Mientras que, la posprueba de escritura fue evaluada mediante rúbrica cuyos criterios, según el rango, son: 4 = capacitado, 3 = se observan logros, 2 = en desarrollo y 1 = en inicio. El promedio máximo para esta prueba es de 4.

En el cuadro 5 se observa que, los estudiantes del GE obtuvieron puntuaciones más altas en la posprueba de lectura y de escritura al compararlas con las puntuaciones obtenidas por los estudiantes del GC. De acuerdo con estos datos, la prueba en la cual los estudiantes de ambos grupos tuvieron mejor ejecución fue la de lectura. No obstante, la prueba de escritura resultó una de mayor dificultad para ambos grupos, aunque el GC resultó con puntuaciones más bajas comparadas con las obtenidas por el GE.

Cuadro 5. Datos obtenidos por los estudiantes participantes de los dos grupos investigados

PARTICIPANTE	*GRUPO	POSPRUEBA	
		Lectura	Escritura
		Valor máximo = 3,00	Valor máximo = 4,00
		$\bar{x}$	$\bar{x}$
1	GC	2,44	1,12
2	GC	2,06	2,31
3	GC	2,33	2,08
4	GC	1,89	1,04
5	GC	1,89	1,38
6	GC	2,22	1,08
7	GC	2,11	1,50
8	GC	2,11	1,00
9	GC	2,33	1,46
10	GC	2,67	2,00
11	GC	2,06	1,27
12	GC	2,33	1,38
13	GC	2,39	1,35
14	GC	2,22	1,77
15	GE	2,67	2,35
16	GE	2,72	2,19
17	GE	2,72	3,65
18	GE	2,72	1,96
19	GE	2,50	2,04
20	GE	2,78	2,04
21	GE	2,56	2,04
22	GE	2,83	1,96
23	GE	2,56	2,88
24	GE	2,78	2,50
25	GE	2,83	2,46
26	GE	2,89	2,27
27	GE	2,67	2,92
28	GE	2,94	3,50
29	GE	2,61	2,08
30	GE	2,72	2,73
31	GE	2,83	2,42

\*Grupo: GC = grupo control; GE = grupo experimental;  $\bar{x}$  = promedio.

Fuente: Elaboración propia.

#### 4.1. Diferencia entre las puntuaciones de la preprueba y de la posprueba de lectura, del GE

El resultado de la prueba Wilcoxon, para determinar si existía diferencia significativa entre las puntuaciones en la pre y la posprueba de lectura en el grupo experimental, resultó en un nivel de significancia de ,000, por lo que se encontró que existe diferencia significativa entre las medianas de las puntuaciones de la preprueba y posprueba de lectura. Además, la prueba  $t$  pareada dio como resultado una diferencia de  $\bar{x} = -,54248$ . El valor de  $t = -9,190$  resultó con una significancia de ,000, por lo que se encontró diferencia significativa entre los promedios de las puntuaciones de la preprueba y posprueba de lectura en estudiantes del grupo experimental. La Figura 2 presenta la comparación de resultados entre pre y posprueba de lectura, del GE.

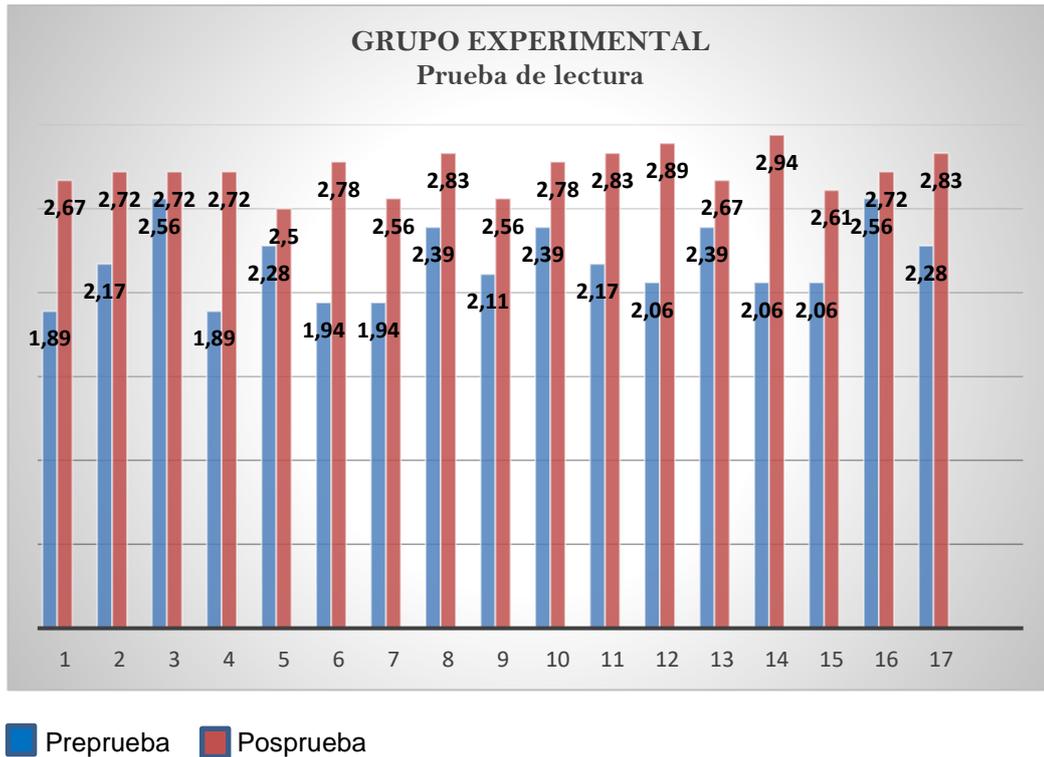


Figura 2. Comparación de resultados entre preprueba y posprueba de lectura, del GE  
Fuente: Elaboración propia.

Se observa en la figura 2 que todos los estudiantes del GE obtuvieron puntuaciones más altas en la posprueba de lectura al compararlas con las obtenidas en la preprueba.

**4.2. Diferencia entre las puntuaciones de la preprueba y de la posprueba de escritura, del GE**

El resultado de la prueba Wilcoxon, para determinar si existe diferencia significativa entre la pre y la posprueba de escritura en el grupo experimental, resultó en un nivel de significancia de ,000, por lo que se encontró que existe diferencia significativa entre las medianas de las puntuaciones de la preprueba y posprueba de escritura. Además, el resultado de la prueba *t* pareada fue de  $\bar{x} = -,54248$ . El valor de  $t = -13,925$  resultó con una significancia de ,000, por lo que se descubrió que existe diferencia significativa entre los promedios de las puntuaciones de la preprueba y posprueba de escritura en estudiantes del grupo experimental. La Figura 3 presenta la diferencia entre pre y pos.

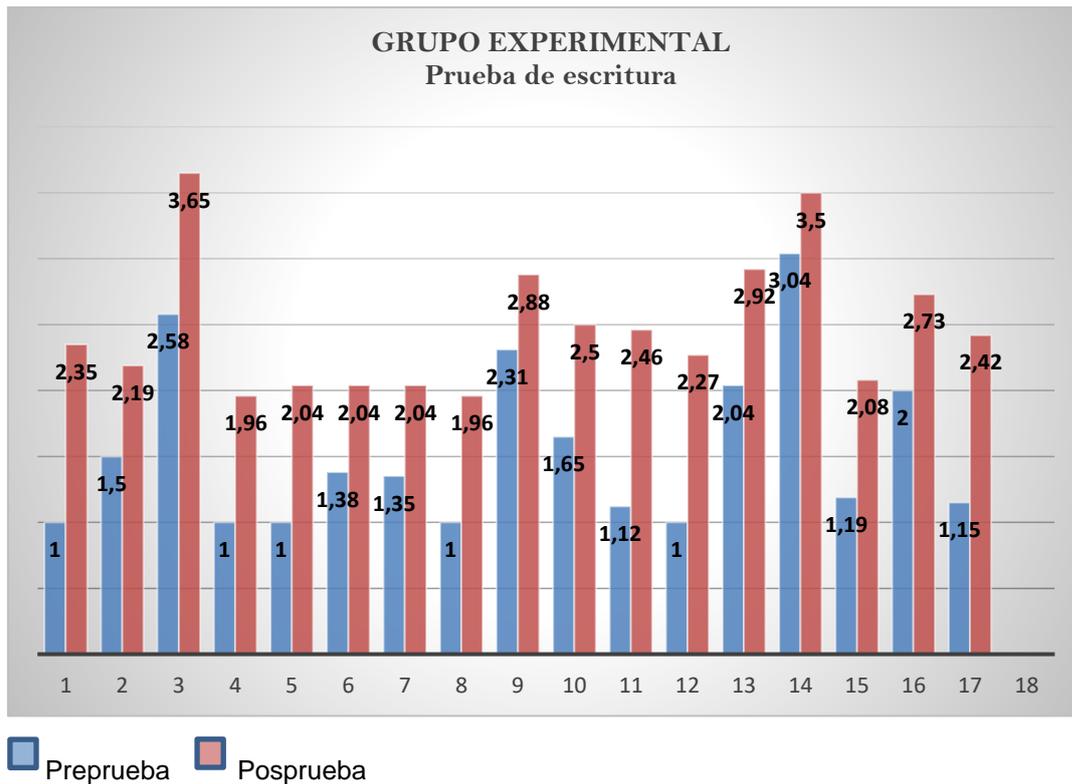


Figura 3. Comparación de resultados entre preprueba y posprueba de escritura, del GE  
 Fuente: Elaboración propia.

Se observa en la figura 3 que todos los estudiantes del GE obtuvieron puntuaciones más altas en la posprueba de escritura al compararlas con las obtenidas en la preprueba.

**4.3. ¿Cuál es el efecto de la integración de las artes visuales al desarrollo de las destrezas de lectura en estudiantes de tercer grado de escuelas públicas del suroeste de Puerto Rico?**

Para contestar esta pregunta y la hipótesis planteada, se llevó a cabo el análisis de regresión lineal simple. Las magnitudes para medir la asociación se clasificaron mediante el uso de la tabla de correlación de Dancy y Reidy (2004). De acuerdo con los resultados, el valor  $r = ,838$  y el valor  $r\ square = ,702$ , lo cual indicó una correlación positiva fuerte ( $0 \leq r \leq 1$ ), entre la preprueba de lectura y las puntuaciones de los estudiantes. Además, el coeficiente de regresión  $B = ,059$  indicó que, por cada punto adicional en la preprueba, la puntuación en la posprueba aumentará en ,059. En este caso, la variable independiente Preprueba de lectura (con un valor de  $t = ,437$ ) tiene un nivel de significancia de ,665, mientras que, la variable independiente Grupo tiene un valor de t positivo ( $t = 10,960$ ) y significancia de ,000. La evidencia apunta que, no hay relación significativa entre las puntuaciones obtenidas por los estudiantes en la preprueba de lectura y en la posprueba de lectura, sin embargo, el estudiante que está en el grupo experimental tendrá mayor probabilidad de obtener puntuaciones más altas en la posprueba de lectura que el estudiante del grupo control, porque el coeficiente de la variable Grupo ( $,506$ ) resultó estadísticamente significativo (significancia = ,000) y positivo.

#### **4.4. ¿Cuál es el efecto de la integración de las artes visuales al desarrollo de las destrezas de escritura en estudiantes de tercer grado de escuelas públicas del suroeste de Puerto Rico?**

Para contestar esta pregunta y la hipótesis planteada, se llevó a cabo el análisis de regresión lineal simple. Las magnitudes para medir la asociación se clasificaron mediante el uso de la tabla de correlación de Dancey y Reidy (2004). De acuerdo con los resultados, el valor  $r = ,958$  y el valor  $r\ square = ,917$ , lo cual indica una correlación positiva fuerte ( $0 \leq r \leq 1$ ) entre la preprueba de escritura y el grupo al que pertenecen los estudiantes con sus puntuaciones en la posprueba. Además, el coeficiente de regresión  $B = ,815$  indica que, por cada punto adicional en la preprueba, la puntuación en la posprueba aumentará en ,815. En este caso, las variables independientes Preprueba de Escritura ( $t = 11,341$ ) y Grupo ( $t = 10,960$ ) son estadísticamente significativas ( $,000 < ,05$ ) y positivas. Hay evidencia estadística de que a mayor puntuación en la preprueba de escritura mayor será la puntuación en la posprueba de escritura. El estudiante que está en el grupo experimental tendrá una mayor probabilidad de obtener más puntuación en la posprueba de escritura que el estudiante del grupo control, porque el coeficiente de la variable Grupo ( $,819$ ) resultó estadísticamente significativo (significancia = ,000), y positivo.

## **5. Discusión y resultados**

El propósito de esta investigación fue determinar el efecto de la integración de las artes visuales en el desarrollo de las destrezas de lectura y escritura en estudiantes de tercer grado en escuelas públicas rurales del suroeste de Puerto Rico. Para la discusión de los resultados, se tomó en consideración la perspectiva de las artes visuales en la educación de Rabkin y Redmond (2004), la teoría de la invención y convención del lenguaje de Goodman (2003) y la teoría del aprendizaje significativo de Ausubel (1962).

El hallazgo de mejoramiento en la lectura coincide con lo que encontró Wilhelm (2002). Este autor realizó una investigación con enfoque cualitativo, a través de un estudio etnográfico cuya duración fue de nueve semanas. En la misma se aplicó el uso de métodos de entrenamiento visual a través de las artes visuales. Los dos estudiantes que participaron en la referida investigación evidenciaron aumento en las destrezas de lectura relacionadas con rapidez, comprensión y aportaciones personales, respecto a la lectura. El hallazgo de esta investigación al cual se hace referencia coincide, además, con los resultados de Rabkin y Redmond (2004b) quienes reportaron que los estudiantes participantes de escuelas cuyo currículo incluía programas de integración de artes visuales en la clase de español, entre otras materias, mejoraron las puntuaciones en las pruebas estandarizadas de lectura más rápidamente que los estudiantes de escuelas que no participaron de dicha integración curricular. Cónsono con los resultados de esta investigación, se encuentran los hallazgos de López-Carrasquillo (2010) quien llevó a cabo un proyecto de alfabetización con adultos que no sabían leer. El investigador reportó que los participantes lograron aprender a leer mediante literatura cuyos temas le interesaban y que eran pertinentes para ellos. Según el autor, este factor contribuyó al desarrollo de la motivación y un mayor compromiso del estudiante con su propio aprendizaje.

Por otro lado, los resultados de incremento en las puntuaciones de la prueba de escritura entre pre y pos, en estudiantes del grupo experimental, coinciden con los resultados que

reportaron Madura (1998), Andrzejczak *et al.* (2005) y Connel-Ross (2006). Los autores identificaron resultados favorables en la escritura, al concluir sus respectivas investigaciones de enfoque cualitativo, a través de las cuales integraron a la escritura aspectos de las artes visuales tales como la utilización y producción de imágenes artísticas y el aspecto estético de la obra. Entre los logros, informaron que la integración de las artes visuales ayudó a los estudiantes a: enfocarse en la tarea, escribir los detalles, desarrollar lenguaje metafórico y sensorial, describir, evaluar e interpretar el tema mediante dibujos y escritos. Así también contribuyó a la motivación de estudiantes reacios a involucrarse en la escritura. Además, el hallazgo al cual se hace referencia coincide con los resultados que presentó Fiske (1999). El autor determinó que los estudiantes que estuvieron expuestos a la educación artística mostraron cambios en el proceso de aprender, tales como, el desarrollo del aprendizaje por descubrimiento y la iniciativa de aprender.

*Efecto de la integración de las artes visuales en lectura.* El efecto positivo del tratamiento en la lectura coincide con los resultados que reportó Catteral (2002) los cuales indican que el entrenamiento visual ayudó a desarrollar la interpretación de textos de los estudiantes que participaron de la integración de las artes visuales. Este resultado es cónsono, además, con Weissman (2004) quien encontró que los estudiantes expuestos a esta variable en una escuela cuyos estudiantes no lograban el desempeño esperado en las pruebas estandarizadas, triplicaron las puntuaciones en el área de comprensión lectora, en el término de cinco años. No obstante, los resultados son contrarios a los de Burger y Winner (2000) ya que los autores no encontraron asociación entre la enseñanza de las artes visuales y el mejoramiento en la lectura.

*Efecto de la integración de las artes visuales en la escritura.* Mediante el análisis de los datos, se determinó que la integración de las artes visuales tuvo un efecto positivo y estadísticamente significativo, tanto en el desarrollo de las destrezas de lectura como en el de las destrezas de escritura. En ambos casos, se encontró que los estudiantes expuestos a la variable artes visuales mostraron un incremento en las puntuaciones obtenidas en las pospruebas con relación a las que habían obtenido en las prepruebas.

Este hallazgo, coincide con el de Deasy (2000), Connel-Ross (2006), Sanders (2006) y Leigh (2007) quienes reportaron que los estudiantes que han estado expuestos a la integración de las artes visuales experimentaron incremento en el aprovechamiento académico, particularmente, en las destrezas de escritura. Además, Eisner (2004) indica que el uso de estrategias de enseñanza que apelan a la motivación y al interés del estudiante tiene como resultado que los aprendices logren desarrollar en mayor escala sus capacidades, comparados con los logros de los estudiantes cuya enseñanza se le ofrece del modo tradicional. Por otra parte, los hallazgos difieren de Craffey (2009) quien luego de realizar una investigación cuasi experimental para determinar la probable relación entre el uso de diversas modalidades artísticas y el desarrollo la adquisición del lenguaje, concluyó que no hubo conexiones significativas entre el uso de las modalidades artísticas y la adquisición del lenguaje.

*Efecto de la integración de las artes visuales en la lectura y la escritura.* El efecto positivo de la integración de las artes visuales tanto en la lectura como en la escritura coincide con la teoría de aprendizaje significativo de Ausubel (1962) quien expuso que este aprendizaje ocurre cuando el individuo incorpora la nueva información de manera que ésta se relacione con la estructura cognitiva previamente existente. Para que esto ocurra, según el autor, el

lector o escritor debe estar motivado, poseer conocimientos previos y construir nuevos significados. En este caso, los estudiantes del grupo experimental estuvieron expuestos a la enseñanza del español como su lengua materna, mientras participaban de manera integrada de actividades artísticas que conectaban con el conocimiento adquirido. Asimismo, en cada experiencia de aprendizaje, el estudiante como lector y escritor seleccionó los conceptos más relevantes y los añadió al nuevo conocimiento lo que hizo del mismo uno más comprensible. De acuerdo con los resultados de la investigación, se infiere que la integración de las artes visuales fue la estrategia que aportó la motivación para lograr los siguientes pasos hacia el desarrollo de las destrezas de lectura y escritura en los estudiantes de tercer grado, del grupo experimental.

En conclusión, la integración de las artes visuales en el desarrollo de las destrezas de lectura y escritura tuvo un efecto positivo en el rendimiento de los estudiantes expuestos a la misma, pues mostraron mejor rendimiento en las pospruebas de lectura y escritura que los estudiantes que no tuvieron esta intervención. No obstante, es importante señalar que estos resultados se limitan a los estudiantes de tercer grado participantes de la investigación, por lo que no son generalizables a la población de estudiantes de tercer grado. Se sugiere para futuras investigaciones que la muestra sea seleccionada aleatoriamente y que se aplique el tratamiento por un período mayor de tiempo.

## Referencias

- Álvarez-Pérez, H. (2006). *Los hallazgos de las neurociencias y su aplicabilidad a la sala de clases: Teoría y práctica*. Guaynabo: Ediciones Santillana.
- Andrzejczak, N., Trainin, G. y Poldberg, M. (2005). From image to text: Using images in the writing process. *International Journal of Education & the Arts*, 6(12).
- Ausubel, D. (1962, November). Learning by discovery. *Educational Leadership*, 20(2), 13-117.
- Ausubel, D. (2002). *Adquisición y retención del conocimiento: Una perspectiva cognitiva*. Barcelona: Paidós.
- Ausubel, D., Novak, J. y Hanesian, H. (2009). *Psicología educativa: Un punto de vista cognoscitivo*. Ciudad de México: Trillas.
- Burger, K. y Winner, E. (Fall 2000). Instruction in visual art: Can it help children learn to read? *Journal of Aesthetic Education*, 34(3-4), 277-293.
- Catterall, J. (2002). The arts and the transfer of learning. En R. Deasy (Ed.), *Critical links: Learning in the arts and student academic and social development* (pp. 151-157). Washington, DC: Arts Education Partnership.
- Coady, K. (2007). *No writer left behind: Examining the reading-writing connection in the reading first classroom through a teacher study group*. Tesis doctoral. Universidad del Estado de Georgia, Atlanta, GA.
- Connel-Ross, B. (2006). *Ekphrasis: Putting the art into language arts drawing on art and visual representation to teach writing*. Tesis doctoral, Universidad de Texas, El Paso, TX.
- Craffey, J. (2009). *An evaluation of how the arts facilitate language acquisition through multisensory modalities*. Tesis doctoral. Universidad de Northcentral, San Diego, CA.
- Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage Publications.

- Cruz, E. y Cruz-López, D. (1979). Un pintor puertorriqueño sobresaliente: Su vida y proyecciones. En A. Pastor, R. Guzmán y K. Hester (Eds.), *Páginas de ayer y de hoy: Por el mundo del cuento y la aventura, libro VI* (pp. 2016-214). River Forest, IL: Laidlaw Brothers.
- Damasio, A. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*. Nueva York, NY: Harvest Book Harcourt.
- Dancey, C. y Reidy, J. (2004). *Statistics without maths for psychology: Using SPSS for Windows*. Londres: Prentice Hall.
- Deasy, R. (2002). *Critical links: Learning in the arts and student academic and social development*. Washington, DC: Arts Education Partnership.
- Departamento de Educación de Puerto Rico. (2013). *Carta circular núm. 5 - 2013 - 2014*. Política pública sobre la organización y la oferta curricular del programa de bellas artes en los niveles elemental y secundario de las escuelas públicas de Puerto Rico. San Juan: Autor.
- Departamento de Educación de Puerto Rico. (2014a). *Carta circular núm. 5 - 2014 - 2015*. Política pública sobre la clasificación de las escuelas e implementación de estrategias de transformación para mejorar el aprovechamiento académico de los estudiantes de acuerdo con el plan de flexibilidad. San Juan: Autor.
- Departamento de Educación de Puerto Rico. (2015a). *Carta circular núm. 4-2015-2016*. Política pública sobre la organización y la oferta curricular del programa de español como lengua materna en los niveles primario y secundario de las escuelas públicas de Puerto Rico. San Juan: Autor.
- Departamento de Educación de Puerto Rico. (2014b). *Estándares de contenido y expectativas de grado de Puerto Rico: Programa de español*. San Juan: Autor.
- Departamento de Educación de Puerto Rico. (2015b). *Informe del perfil escolar*. San Juan: Autor.
- Departamento de Educación de Puerto Rico. (2003a). *Marco curricular: Programa de bellas artes*. San Juan: Instituto Nacional para el Desarrollo Curricular.
- Departamento de Educación de Puerto Rico. (2003b). *Marco curricular: Programa de español*. San Juan: Instituto Nacional para el Desarrollo Curricular.
- Dewey, J. (1934). *Art as experience*. Nueva York, NY: Perigree Books.
- Dollins, C. (2001). *The connection between reading and writing for early readers*. Tesis doctoral, Universidad de Pepperdine, Malibu, CA.
- Eisner, E. (1972). *Educar la visión artística*. Barcelona: Paidós Ibérica.
- Eisner, E. (2004). *El arte y la creación de la mente*. Barcelona: Paidós Ibérica.
- Fiske, E. (1999). *Champions of change: The impact of the arts of learning*. Washington, DC: Arts Education Partnership.
- Fons, M. (2010). *Leer y escribir: 10 ideas claves para los primeros pasos. Con firma 2010: Leer para aprender. Leer en la era digital*. Recuperado de <https://sede.educacion.gob.es/publiventa/con-firma-2010-leer-para-aprender-leer-en-la-era-digital/educacion/13939>. <https://doi.org/10.4438/978/-84-369-4970-4>
- Freire, P. (2004). *Cartas a quien pretende enseñar*. Buenos Aires: Siglo XXI.
- Goodman, K. (2003). El aprendizaje y la enseñanza de la lectura y la escritura. *Enunciación*, 8(1), 77-78.

- Goodman, K. (2002). El proceso de lectura: Consideraciones a través de las lenguas y del desarrollo. En E. Ferreiro y M. Palacios, M. (Eds.), *Nuevas perspectivas sobre los procesos de lectura y escritura* (pp. 13-28). Coyoacán: Siglo Veintiuno Editores.
- Goodman, K. (1989). *Lenguaje integral*. Mérida: Editorial Venezolana.
- Goodman, K. (1995). Lenguaje integral: Un camino fácil para el desarrollo del lenguaje. *Textos de didáctica de la lengua y la literatura*, 3(3), 77-91.
- Goodman, Y. y Altwerger, B. (1981). *A study of the development of literacy in preschool children. Occasional Papers, Program in Language of Literacy*. Tucson, AZ: University of Arizona.
- Graves, D. (1973). *Children's writing: Research directions and hypotheses based upon an examination of the writing processes of seven years old children*. Tesis doctoral. Universidad de Nueva York, Buffalo, NY.
- Hartse, J., Burke, C. y Woodward, V. (1982). Children's language and world: Initial encounters with print. En J. Langer y M. Smith-Burke (Eds.), *Bridging the gap: Reader meets author* (pp. 105-131). Newark, DE: International Reading Association.
- Leigh, S. (2007). *Drawing, writing, and second graders*. Tesis doctoral. Universidad del Sur de California, Columbia, SC.
- Kay, L. (2008). *Art education pedagogy and practice with adolescent students at -risk in alternative schools*. Tesis doctoral, Universidad del Norte de Illinois, DeKalb, IL.
- Kontos, S. y Huba, N. (1983, junio). The development and function of print awareness. Comunicación presentada en el *Biennial Meeting of the Society for Research in Child Development*. ERIC Document Reproduction Service No. ED 230-237.
- Krishna, P. (2004). *La educación correcta para el siglo XXI*. Recuperado de [http://www.pkrishna.org/spanish/right\\_education\\_spanish.html](http://www.pkrishna.org/spanish/right_education_spanish.html)
- López-Carrasquillo, M. (2010). *Marginados en un mundo de letras: 14 principios para enseñar con amor, compasión y alegría*. San Juan: Reliable Graphics.
- Madura, S. (1998). *Transitional readers and writers respond to literature through discussion, writing and art*. Tesis doctoral, Universidad de Nevada, Reno, NV.
- Maeso-Rubio, F. (2003). El arte de construir el conocimiento artístico. El diseño curricular de educación artística en la escuela primaria. En R. Marín Viadel (Ed.), *Didáctica de la educación artística* (pp. 229-271). Madrid: Pearson Educación.
- Mass, L. (1982). Developing concepts of literacy in young children. *The Reading Teacher*, 35(6), 670-675.
- McMillan, J. y Schumacher, S. (2010). *Research in education: Evidence-based inquiry*. Boston, MA: Pearson Education.
- Public Law 89-10. (1965, April). *Elementary and Secondary Education Act*. Washington, DC: United State Congress.
- Public Law 107-110. (2002, January). *An act to close the achievement gap with accountability, flexibility, and choice, so that no child left behind*. Washington, DC: 107th United State Congress.
- Rabkin, N. y Redmond, R. (2004a). Learning and the arts. En N. Rabkin y R. Redmond (Eds.), *Putting the arts in the picture: Reframing education in the 21st century* (pp. 5-15). Chicago, IL: Columbia College Chicago.

- Rabkin, N. y Redmond, R. (2004b). Putting the arts in the picture: Reframing education in the in the 21st century. En N. Rabkin y R. Redmond (Eds.), *Putting the arts in the picture: Reframing education in the 21st century* (pp. 127-157). Chicago, IL: Columbia College Chicago.
- Rabkin, N. y Redmond, R. (2006). The arts make a difference. *Educational Leadership*, 63(5), 60-64.
- Rachlin, C. (2008). *I have something to say: Using art to teach the writing of persuasive essays*. Tesis de Maestría. Universidad de California, San Diego, CA.
- Reyes-Tosta, M. (2013). El aprendizaje del español y del inglés en Puerto Rico y sus influencias extranjeras. *Scientific International Journal*, 10(3), 44-53.
- Rolón-Collazo, L. (2013). Leer para escribir y escribir para leer: Propuesta sobre el proceso de escritura, el entorno como texto y la complicidad entre lectura y escritura. *El Sol: Revista de la Asociación de Maestros de Puerto Rico*, 54(2), 32-34.
- Ruiz, D. (1998). *La alfabetización temprana en el ambiente preescolar: Nuevas perspectivas para aprender a leer y escribir*. Río Piedras, PR: Departamento de Programas y Enseñanza, Universidad de Puerto Rico.
- Salkind, N. (1999). *Métodos de investigación*. Ciudad de México: Prentice Hall.
- Sanders, J. (2006). *Qualitative study of students' composing experiences in art-infused writing curriculum*. Tesis doctoral. Universidad de Florida, Gainesville, FL.
- Santiago-Villafañe, R. y Figarella-García, F. (2012). Más allá de las disciplinas: Competencia para la vida. *El Sol: Revista de la Asociación de Maestros de Puerto Rico*, 53(3), 22-26.
- Senitt, A. (2006). *Signs of substance: An investigation regarding the confluence of art and language in an elementary school context*. Tesis doctoral. Universidad Concordia, Canadá.
- Schneuwly, B. (2010). *¿Qué es escribir? Herramientas y aprendizaje de la escritura. Con firma 2010: Leer para aprender. Leer en la era digital*. Madrid: Ministerio de Educación. <https://doi.org/10.4438/978/-84-369-4970-4>
- Solé, I. (2010). *Ocho preguntas en torno a la lectura y ocho respuestas no tan evidentes. Con firma 2010: Leer para aprender. Leer en la era digital*. Madrid: Ministerio de Educación. <https://doi.org/10.4438/978/-84-369-4970-4>
- Sousa, D. (2002). *Como aprende el cerebro: Una guía para el maestro en la clase*. Thousands Oaks, CA: Corwin Press.
- United States Department of Education. (2015). *Puerto Rico Elementary & Secondary Education Act Flexibility Request: Approved*. Washington, DC: Author.
- Vygotsky, L. (1995). *Pensamiento y lenguaje*. Buenos Aires: Ediciones Fausto.
- Weissman, D. (2004). You can't get much better than that. En N. Rabkin y R. Redmond (Eds.), *Putting the arts in the picture: Reframing education in the 21st century* (pp. 17-48). Chicago, IL: Columbia College Chicago.
- Wilhelm, J. (2000). Reading is seeing: Using visual response to improve the literary Reading of reluctant readers. En R. Deasy (Ed.), *Critical links: Learning in the arts and student academic and social development* (pp. 48-72). Washington, DC: Arts Education Partnership.

## **Breve CV de los autores**

### **Wanda I. Padilla Martínez**

Profesora de Artes Visuales del Departamento de Educación de Puerto Rico. Ha coordinado proyectos para estudiantes, dirigidos a fortalecer las destrezas académicas básicas; entre éstos: Ayúdame a estudiar, Exploradores del ambiente y Pequeños Periodistas. Además, perteneció a las organizaciones Academia de Mentores de Puerto Rico, Coaching for Change y Living with a Star: Starpartners, Capítulo de Puerto Rico, NASA, entre otras. Doctora en Educación: Currículo y Enseñanza de la Pontificia Universidad Católica de Puerto Rico, Recinto de Ponce. Maestría en Artes Plásticas: Grabado, de la Universidad Interamericana de Puerto Rico, Recinto de San Germán y Bachillerato en Educación Secundaria en Arte de la Universidad de Puerto Rico, Recinto de Río Piedras. ORCID ID: 0000-0002-2696-2983. Email: padilla.wanda@gmail.com

### **Jaime Ortiz-Vega**

Catedrático y Director de la Escuela Graduada en Educación de la Pontificia Universidad Católica de Puerto Rico. Doctor en Educación: Gerencia y Liderazgo Educativo de New York University. Consultor para agencias gubernamentales y privadas en áreas de planificación estratégica, gerencia educativa, currículo y enseñanza. Miembro de la American Educational Research Association y de Association for Supervision and Curriculum Development, International Level. Ex presidente del Capítulo de Puerto Rico de Asociación for Supervision and Curriculum Development. Poseedor de varias publicaciones en revistas académicas. ORCID ID: 0000-0003-0310-9362. Email: jortiz@pucpr.edu



# Influencia de la Longitud, la Complejidad y la Inferencia en la Resolución de Pruebas de Comprensión Lectora Inferencial

## Effects of Length, Complexity, and Inference in the Resolution of Inferential Reading Comprehension Tests

Alejandra Platas-García \*<sup>1 2</sup>

J. Martín Castro-Manzano <sup>2</sup>

Verónica Reyes-Meza <sup>3</sup>

Irene Gaona-Gordillo <sup>3</sup>

<sup>1</sup>Benemérita Universidad Autónoma de Puebla (BUAP), México

<sup>2</sup>Universidad Popular Autónoma del Estado de Puebla (UPAEP), México

<sup>3</sup>Universidad Autónoma de Tlaxcala (UATx), México

Este trabajo surge a partir de nuestra experiencia docente a nivel superior con las tareas de revisión y elaboración de pruebas de opción múltiple para medir comprensión lectora. Nuestros objetivos son dos: 1) determinar si las variables de longitud, complejidad e inferencia de una prueba de comprensión lectora inferencial tienen efectos en el desempeño de los estudiantes y, consecuentemente, 2) argumentar a favor de la elaboración de pruebas que promuevan o incluyan el uso de inferencias no-deductivas en adición a las deductivas. Para lograr estos objetivos llevamos a cabo un estudio con 187 estudiantes universitarios a los que les aplicamos una prueba de comprensión lectora inferencial de opción múltiple. Los datos obtenidos fueron analizados con un modelo lineal generalizado mixto. Entre los resultados generales encontramos que los estudiantes tienen un mejor desempeño en la prueba de comprensión lectora cuando se les solicita realizar inferencias no-deductivas y los textos son complejos y largos; mientras que su desempeño decrece cuando se les solicita realizar inferencias deductivas y los textos son simples y cortos. Estos resultados parecen soportar la tesis de que existe un efecto de las características de las pruebas en la comprensión lectora de los estudiantes, lo cual, a su vez, nos permite sugerir que el diseño de pruebas de comprensión lectora podría enriquecerse tomando en cuenta un concepto más incluyente de inferencia que sería consistente con el concepto de racionalidad acotada.

**Palabras clave:** Evaluación de estudiantes; Razonamiento; Deducción; Tests de lectura; Tests de elección múltiple; Comprensión lectora; Inferencias.

Our interest in this research comes from our experience at reviewing and developing multiple-choice tests in order to measure reading comprehension skills at an undergraduate level. We have two goals: 1) to determine if the variables of length, complexity, and inference, of a reading comprehension test have an effect over students' performance and, consequently, 2) to argue in favour of tests that promote or include non-deductive inference in addition to deductive inference as a parameter of their design. To reach these goals we have applied an inferential reading comprehension test to a group of 187 undergraduate students. The obtained data was analysed with a generalized linear mixed model. Among the general results, we find that students perform better when the test requires them to carry out a non-deductive inference over complex and long texts; whereas their performance decreases when the test calls for a deductive inference over simple and short texts. These results seemingly support the claim that there is an influence of the properties of the test over the students reading comprehension, which, in turn, allows us to suggest that the design of reading comprehension tests might be enriched by taking into account a more inclusive concept of inference consistent with the notion of bounded rationality.

**Keywords:** Student evaluation; Reasoning; Deduction; Reading tests; Multiple choice tests; Reading comprehension; Inferences.

---

\*Contacto: [aplatasg@gmail.com](mailto:aplatasg@gmail.com)

issn: 1989-0397

[www.rinace.net/riee/](http://www.rinace.net/riee/)

<https://revistas.uam.es/riee>

Recibido: 24 de noviembre de 2017

1ª Evaluación: 18 de diciembre de 2017

Aceptado: 30 de enero de 2018

## 1. Introducción

Este trabajo surge a partir de nuestra experiencia docente a nivel superior con las tareas de revisión y elaboración de pruebas de comprensión lectora, específicamente, pruebas de opción múltiple en las que se presenta un texto a partir del cual se realizan preguntas y se ofrecen, comúnmente, tres o cuatro opciones de respuesta y sólo una de ellas es correcta. En el contexto de este tipo de pruebas creemos que es pertinente reconsiderar las características de su diseño, porque podría existir una influencia del mismo en el desempeño de los estudiantes; lo que, de ser cierto, nos podría ofrecer elementos de juicio para argumentar a favor de la elaboración de pruebas que promuevan una comprensión inferencial enriquecida, dado que, por ejemplo, como afirman Guerra y Guevara (2013), los problemas de comprensión lectora en estudiantes universitarios residen en la dificultad de identificar “la o las ideas principales y [...] el uso apropiado de la información para hacer inferencias y realizar un análisis crítico sobre la base de lo leído” (p. 278).

Motivados por estas consideraciones, hemos analizado ejemplos de pruebas de comprensión lectora de opción múltiple con el propósito de revisar algunas de sus propiedades estructurales: características materiales que comparten (longitud), cualidades que pueden ser distinguidas (complejidad) y el tipo de inferencia que propician (deductiva, no-deductiva). De acuerdo al modelo de van Dijk (1992), las propiedades estructurales de un texto son tres: 1) la *microestructura* que engloba los enunciados y las secuencias de enunciados, 2) la *macroestructura* que proporciona la información que dirigirá la interpretación de las palabras y los enunciados, y que representa la coherencia global del texto, y 3) la *superestructura* que es un tipo de esquema abstracto que establece el orden general al cual el texto se adapta. Dado este modelo, en esta colaboración trabajamos una analogía entre los atributos estructurales de los textos y las propiedades estructurales de las pruebas previamente mencionadas: 1) la microestructura se asocia con la longitud, 2) la macroestructura con la Complejidad, y 3) la superestructura con los tipos de inferencia.

Así pues, este estudio tiene dos objetivos: 1) conocer si la longitud (corto, largo), la complejidad (simple, complejo) y el tipo de inferencia (deductiva, no-deductiva) de una prueba de comprensión lectora inferencial influyen en el desempeño de los estudiantes y, consecuentemente y 2) argumentar a favor de la elaboración de pruebas que promuevan inferencias no-deductivas además de las deductivas. La organización del trabajo es la siguiente. En la primera sección ofrecemos una breve caracterización del concepto de comprensión lectora que asumimos como marco teórico general. Posteriormente, en la segunda, describimos las propiedades estructurales de longitud, complejidad e inferencia. En la tercera, analizamos una prueba de comprensión lectora a la luz de sus características de superestructura. En la cuarta, exponemos el concepto de inferencia extendida y racionalidad acotada. En la quinta, mostramos la metodología que seguimos. En la sexta, los resultados obtenidos. Finalmente, en la séptima, presentamos una breve discusión y algunas conclusiones.

## 2. Comprensión lectora

Existen tres enfoques principales sobre el proceso de comprensión lectora, a saber: 1) el *modelo abajo-arriba (bottom up)* que considera que el proceso de comprensión procede de la percepción visual, el reconocimiento de letras y palabras, la agrupación de palabras para

formar proposiciones y así sucesivamente, cada operación es requerida para que se realice la siguiente; 2) el *modelo arriba-abajo (top down)* para el que los aspectos perceptivo-visuales están subordinados a los conocimientos semánticos y sintácticos del lector; y 3) los *modelos interactivos* para los que el proceso de comprensión está dirigido en forma interactiva tanto por el texto como por el conocimiento del lector (Acosta, 2009). En esta investigación, la comprensión lectora es entendida bajo el tercer enfoque por lo que se define como “un proceso a través del cual el lector elabora un significado en su interacción con el texto” (Pérez Zorrilla, 2005, p. 123; cf. Guerra y Guevara, 2017, p. 79).

En el proceso de comprensión lectora, el lector aporta sus esquemas de conocimiento (que son fruto tanto de sus experiencias como de aprendizajes previos) con el fin de integrar los datos nuevos incluidos en el texto; en este proceso, los esquemas del lector pueden sufrir modificaciones y enriquecimientos continuos (Solé, 1987). (Sobre los esquemas véase: Anderson y Pearson, 1987; Intraversato, 2013; Salazar, 2006).

Como se puede inferir, el proceso de comprensión lectora requiere que el lector posea un conjunto articulado de conocimientos y habilidades que van más allá de una simple decodificación o comprensión literal del texto (esto es consistente con el modelo de García, 1993, que propone niveles ordenados: decodificación, comprensión literal, comprensión inferencial y metacompreensión), por esto, es un proceso activo e interactivo de construcción de significados en el que están implicados el lector, el texto y el contexto (Giovannini y Rosa, 2015).

En este trabajo asumimos un nivel de comprensión lectora de nivel inferencial siguiendo la clasificación de García (1993), para quien la elaboración de inferencias forma parte del procesamiento textual, el cual está conformado por representaciones, por la integración de enunciados, el conocimiento del mundo por parte de un lector (lo cual es consistente con el concepto de conocimientos enciclopédicos en la propuesta de Intraversato, 2013), “las inferencias según los esquemas cognitivos que el mismo lector posee y su interpretación del texto” (García, 1993, p. 94).

Por estas razones la comprensión lectora representa una tarea fundamental en la formación académica y, por tanto, en el contexto educativo; por ello los docentes y las instituciones educativas se interesan por conocer cuál es el nivel de comprensión lectora que poseen los estudiantes. Sin embargo, la medición de este nivel de comprensión no es transparente, por lo que para obtener evidencias es necesario realizar evaluaciones indirectas (Consejo de Europa, 2002), ya que como afirma Pérez Zorrilla (2005), “dado que no se puede observar directamente la comprensión lectora, hay que pedirle al alumno que realice algún tipo de tarea que indique el grado de comprensión alcanzado” (p. 128). Esta evaluación indirecta se suele llevar a cabo mediante pruebas en las que se solicita al estudiante que realice tareas como marcar casillas, completar frases, contestar preguntas o elegir una opción correcta de entre una serie finita de alternativas. Dentro de este conjunto de pruebas, una que es utilizada frecuentemente para evaluar la comprensión lectora es la prueba de opción múltiple. En la siguiente sección revisamos la estructura de este tipo de pruebas trabajando la analogía con el modelo de van Dijk.

### 3. Propiedades estructurales de las pruebas

Para estudiar los atributos de micro (longitud), macro (complejidad) y superestructura (inferencia) hemos diseñado una prueba de comprensión lectora inferencial de opción múltiple que usamos como instrumento para esta investigación. En la sección de metodología ofrecemos más detalles de la prueba.

#### 3.1 Longitud

Si bien esta característica es importante para definir la complejidad de un texto, en este estudio la hemos distinguido de la complejidad *per se* por razones que mencionaremos en el siguiente apartado. Para definir la longitud del texto se pueden emplear distintos criterios, por ejemplo: el número de palabras (cf. Rice, 1994, para quien un texto corto es de 142 palabras; mientras que para Bell, 2001, un texto corto es de 300) o el número de páginas (cf. Mangen, Walgermo y Brønning, 2013, para quienes cuatro páginas forman un texto corto). El criterio empleado en esta investigación fue el número de líneas que tiene cada texto asumiendo que: 1) los márgenes y el tamaño de las fuentes son fijos e invariables y 2) la información en los textos no es redundante. De esta forma, en esta investigación, un texto es *corto*, si el número de líneas es menor o igual a 4; y es *largo*, si el número de líneas es mayor o igual a 5 y menor o igual a 8. Este es un criterio viable para el análisis de las características del texto presentado, ya que si los textos fueran de una extensión mayor a los rangos mencionados, harían de la prueba un instrumento inviable en términos del tiempo utilizado para su aplicación.

#### 3.2 Complejidad

De acuerdo con Ferrari (2003), para conocer la dificultad de un texto existen dos criterios, uno cuantitativo, la *legibilidad*, y otro cualitativo, la *comprensibilidad*. Con respecto al primero, la autora afirma que las investigaciones sobre legibilidad tienen como objetivo la construcción de funciones matemáticas que tomen como entrada un texto y asignen un índice numérico que ilustre su dificultad, considerando como factores de complejidad las frases largas y las palabras poco comunes, por ejemplo; sin embargo, estas funciones no necesariamente toman en cuenta las competencias y conocimientos individuales del lector; del mismo modo, excluyen otros aspectos importantes de un texto, como la estructura lógica o la presencia de tablas y figuras. Por esto último, Ferrari (2003) sugiere que la legibilidad no puede considerarse como un índice preciso de la dificultad de un texto.

Además, sobre la misma legibilidad, Pallotti (2013) afirma que este tipo de funciones matemáticas se basan en el cálculo de la longitud media de los enunciados y de las palabras: textos con enunciados largos y palabras largas tendrán un valor de legibilidad más bajo que los textos con enunciados cortos y palabras cortas. Sin embargo, este indicador no basta para afirmar que un texto es simple o complejo, puesto que es necesario acudir al segundo criterio, la comprensibilidad, ya que un texto que contiene únicamente enunciados cortos y palabras cortas puede resultar obscuro o poco asequible debido a posibles lagunas informativas.

Por esto, el segundo criterio, la comprensibilidad, parece ser más útil para el análisis de la dificultad de un texto, pues los estudios de comprensibilidad no tienen por objetivo obtener índices numéricos, sino explorar y utilizar metodologías de análisis de dificultad de un texto como las que ofrece la lingüística textual, que ha identificado dos propiedades

fundamentales de un texto, a saber: la cohesión y la coherencia. Entendiéndose por *cohesión* la serie de mecanismos que aseguran la unión entre las partes de los textos a nivel superficial, es decir, el respeto de las relaciones formales entre los enunciados (coordinación y subordinación), las referencias temporales, el orden de las palabras y el uso de conectores como adverbios o conjunciones que sirven para unir explícitamente dos enunciados; la *coherencia*, por otro lado, está relacionada con los procesos lógicos y conceptuales que permiten la producción de sentido (Ferrari, 2003). En suma, mientras la legibilidad es una característica intrínseca del texto y es medible, la comprensibilidad es una característica proporcional, pues depende de la distancia de los conocimientos y competencias del lector al texto, a su estructura, a su contenido (Zambelli, 2014).

Pues bien, dado que la clasificación de un texto como simple o complejo no es una tarea sencilla y no basta con una métrica de legibilidad o longitud, optamos por tomar como criterio de complejidad el juicio de un grupo de expertos asumiendo la noción de comprensibilidad y tomando en cuenta los siguientes criterios: verbos, vocabulario, redacción, extensión y tema. De esta forma, dado un universo de 7 lectores expertos (escritores, editores, docentes de español para extranjeros), un texto es *simple*, si 4 o más expertos etiquetan un texto como “simple”; un texto es *complejo*, si 4 o más expertos lo etiquetan como “complejo”.

### 3.3 Inferencia

Pallotti (2013) afirma que en el proceso activo de la comprensión de un texto el lector busca darle sentido a lo que lee revisando la coherencia entre el contenido del texto y sus conocimientos previos; además, dado que un texto no puede decirlo todo y mucha información se da por sobreentendida, el lector busca unidad a través de una serie de inferencias con las cuales llena los vacíos e integra la información en un todo coherente; del mismo modo, extrae las consecuencias de lo que está leyendo, usando ocasionalmente su conocimiento general (p. 3). La inferencia, por tanto, es un proceso relacionado con la comprensión lectora en el que a partir de la aceptación de ciertas información (premisas) se llega a la aceptación de cierta información (conclusión) (Copi y Cohen, 2007; Cook, 2009).

Una inferencia puede ser de tres distintos tipos, dependiendo de la relación entre las premisas y la conclusión. En algunas inferencias la conclusión no añade información nueva a lo que se ha expresado en las premisas; mientras que en otros tipos de inferencia la conclusión sí puede añadir información nueva. Esta distinción básica induce la división operacional entre inferencias deductivas y no-deductivas (inducción y abducción) que utilizamos para este estudio.

- a) *Deducción*. En sentido estricto, una *deducción* se define como una secuencia finita de enunciados dentro de un sistema formal donde cada enunciado en la secuencia es un axioma, una suposición o el resultado de la aplicación de una regla de inferencia a uno o varios enunciados precedentes. El enunciado final es la conclusión del argumento que ha sido derivado y cada suposición es una premisa del argumento derivado (Cook, 2009). De manera informal, en una deducción la información está completa y por ende la conclusión no agrega nada nuevo a lo que ya está expresado en las premisas (Rodríguez Rodríguez, 2005). Por ello, en una inferencia deductiva es imposible que las premisas sean verdaderas y la conclusión sea falsa (Cook, 2009). Un ejemplo típico de inferencia deductiva sería el siguiente:

1. *Todos los hombres son mortales.*

2. *Sócrates es un hombre.*

C. *Por tanto, Sócrates es mortal.*

- b) *Inducción.* El concepto más difundido de inducción afirma que esta es una inferencia que va de lo particular a lo general. Como este concepto es muy ambiguo, se ha desarrollado un concepto más claro: una inducción es una inferencia que permite extraer conclusiones a partir de información incompleta dado cierto grado de soporte (Hawthorne, 2014). El grado de soporte que proveen las premisas ofrece la base conceptual que sostiene a la conclusión, pero este grado de soporte no es deductivo, no implica necesidad, sino probabilidad; por ello, a diferencia de la inferencia deductiva, la inductiva bien podría partir de premisas verdaderas y llegar a conclusiones falsas (Cook, 2009). Consideremos un ejemplo de inferencia inductiva adaptado de Okasha (2002):

1. *Los cinco primeros huevos de esta caja estaban descompuestos.*

2. *Los huevos en la caja tienen la misma fecha de caducidad.*

C. *Por tanto, muy probablemente, el sexto huevo estará descompuesto también.*

- c) *Abducción.* Una abducción es una inferencia cuyo producto es una explicación, con cierta estructura inferencial, que se dispara por un hecho que necesita explicarse (Aliseda, 1998, p. 10-11). Como la inducción, la abducción no requiere criterios de necesidad sino de probabilidad y es también un tipo de inferencia sintética. Según el planteamiento de Aliseda (1997, 1998), la diferencia entre inducción y abducción consiste en que mientras la última es una inferencia que parte de un hecho particular para buscar su explicación, la primera parte de una serie de observaciones para llegar a enunciados generales. A continuación mostramos un ejemplo de abducción adaptado de Okasha (2002):

1. *El queso en la alacena desapareció, excepto por unas pocas migajas.*

2. *Anoche se oyeron ruidos que provenían de la alacena.*

C. *Luego, seguramente, un ratón se comió el queso.*

Además de las propiedades estructurales previamente descritas, es importante hacer el siguiente par de aclaraciones antes de continuar con nuestra exposición. Primero, es posible clasificar los textos como continuos o no-continuos por su estructura interna y el procesamiento requerido para su comprensión (Serrano, Vidal-Abarca y Ferrer, 2017); y segundo, es posible dividirlos en narrativos, descriptivos, expositivos y argumentativos de acuerdo a su tipología textual (Sánchez Lobato, 2007). Dadas estas consideraciones, para el diseño de nuestra prueba hemos elegido trabajar con textos continuos-descriptivos. Continuos, porque este tipo de textos presentan un orden secuencial determinado que ayuda a los estudiantes a entender las ideas “relacionándolas unas con otras mediante inferencias, hasta formar una representación mental coherente” (Serrano et al., 2017, p. 283); descriptivos, porque “presentan con claridad y rigor los rasgos característicos de personas, animales, objetos, lugares, fenómenos o situaciones” (Sánchez Lobato, 2007, p. 61).

## 4. Análisis de una prueba de comprensión lectora a la luz de los tipos de inferencia

Si bien las propiedades estructurales de longitud y complejidad no son triviales, parece que son más fáciles de identificar, a diferencia de la propiedad de inferencia. Por ello, en esta sección ilustramos la importancia de reconsiderar la propiedad de inferencia a través del análisis de la sección de comprensión lectora del examen *Diplomas de Español como Lengua Extranjera* (DELE) en la prueba de opción múltiple a partir de los ejemplos disponibles en el sitio web del Instituto Cervantes para los niveles del Marco común europeo de referencia para las lenguas (MCERL): A1, A2, B1, B2, C1. El DELE es un examen que acredita el grado de competencia y dominio del idioma español y que otorga el Ministerio de Educación, Cultura y Deporte de España. El Instituto Cervantes se encarga de organizar las convocatorias de los exámenes, mientras que la Universidad de Salamanca, en España, se dedica a la elaboración, corrección y evaluación final de los exámenes. En el cuadro 1 presentamos algunos casos representativos que permiten apreciar el tipo de inferencia que solicitan las versiones consultadas.

Cuadro 1. Ejemplos de inferencias solicitadas en una prueba de comprensión de lectura del examen DELE

DATOS DEL EXAMEN DELE	FRAGMENTO DEL TEXTO PARA RESPONDER A LA PREGUNTA	PREGUNTA Y OPCIONES DE RESPUESTA	COMENTARIOS SOBRE EL TIPO DE INFERENCIA
Nivel A1 Tarea 1 20 de mayo de 2011	<p>¡Hola María!</p> <p><i>Te escribo desde Málaga. Mi empresa necesita un informe sobre el turismo en el sur de España y voy a pasar un mes aquí.</i> Es un trabajo difícil porque no hablo muy bien español y utilizo mucho el diccionario... Volveré a Francia el día treinta de este mes... Por las mañanas, me levanto a las ocho, desayuno en el hotel y, después, salgo a trabajar. Cada día visito una parte diferente de la costa..."</p>	<p>1. Pierre escribe un correo electrónico sobre...</p> <p>a) sus vacaciones en la playa. b) su trabajo en Francia. c) su curso de español en Málaga. d) su trabajo en el sur de España.</p>	<p>La respuesta correcta es la opción d), como muestra la siguiente deducción:</p> <p>1. Si Pierre está escribiendo un informe, está trabajando. 2. Pierre está escribiendo un informe. 3. Por tanto, Pierre está trabajando. 4. Si Pierre está en Málaga, está en el sur de España. 5. Pierre está en Málaga. 6. Pierre está en el sur de España. 7. Por tanto, Pierre está trabajando en el sur de España.</p>
Nivel A2 Tarea 2 20 de noviembre de 2010	<p>"Hace ya días que quiero hacer una cena para enseñaros mi casa a todos los del grupo. Este fin de semana estoy muy ocupada, pero <i>he decidido hacerla el próximo martes</i> a las 21:30. ¿Puedes venir? Había pensado hacerla en el jardín, pero parece que</p>	<p>11. La cena será...</p> <p>a) este fin de semana. b) la próxima semana. c) en el jardín.</p>	<p>Pregunta 11. La respuesta correcta es la opción b), como muestra la siguiente deducción:</p> <p>1. El "próximo martes" implica la "próxima semana". 2. La cena será el "próximo martes".</p>

<p><b>Nivel B1</b> <b>Tarea 2</b> <b>23 de agosto de 2013</b></p>	<p>va a llover, así que mejor cenamos en el salón.”</p> <p>“Ahora Ana trabaja menos, por eso puede hacer una de las cosas que más le gustan: <i>ir al gimnasio muy temprano</i>; también le interesa ir al cine, le encanta ver exposiciones de pintura y, lo más importante, pasar su tiempo libre con su familia y sus amigos.”</p>	<p>12. A Ana le gusta...</p> <p>a) hacer deporte por la mañana. b) pintar en su tiempo libre. c) ir al cine con su familia.</p>	<p>C. La cena será la “próxima semana”.</p> <p>Pregunta 12 La respuesta correcta es la opción a), como muestra la siguiente deducción:</p> <p>1. Si a Ana le gusta ir al gimnasio muy temprano, a Ana le gusta hacer deporte por la mañana. 2. A Ana le gusta ir al gimnasio muy temprano. C. A Ana le gusta hacer deporte por la mañana.</p>
<p><b>Nivel B2</b> <b>Tarea 1</b> <b>23 de agosto de 2013</b></p>	<p>“Fue una experiencia divertida en la que se mezclaban varias cosas: el compañerismo, la competitividad y la extraña sensación de poder realizar esta actividad como parte del trabajo”, recuerda Eduardo Loyola, directivo de la consultoría Interface. Loyola fue uno de los conejillos de Indias de un programa diseñado por Óscar de las Mozas, coautor del estudio. “Entrenábamos fuera del horario laboral. Al principio, a algunos compañeros les resultó complicado (sentían vergüenza de verse con indumentaria deportiva), pero pronto <i>nos encontramos trotando por el monte o por pistas de atletismo</i> sin ningún problema. Creo que tener una hora de entrenamiento con los compañeros favorece el trabajo en equipo”, dice Loyola.</p>	<p>3. En el texto Eduardo Loyola dice que...</p> <p>a) se sentía raro haciendo deporte. b) el programa incluía deporte al aire libre. c) gracias al entrenamiento rendía más.</p>	<p>Pregunta 3 La respuesta correcta es la opción b), como muestra la siguiente deducción:</p> <p>1. Si Eduardo Loyola trota por el monte o por pistas de atletismo, el programa incluye deporte al aire libre. 2. Eduardo Loyola trota por el monte o por pistas de atletismo. C. El programa incluye deporte al aire libre.</p>
<p><b>Nivel C1</b> <b>Tarea 1</b> <b>19 de noviembre de 2011</b></p>	<p>“Teleplán. ... En cumplimiento de la normativa vigente nos ponemos en contacto con usted para reiterarle los derechos que le asisten como usuario de nuestros servicios. Tales derechos, detallados a continuación, le serán comunicados</p>	<p>1. Según este documento, el cliente de Teleplán...</p> <p>a) puede descargar su contrato desde la página web de la compañía. b) recibirá comunicación</p>	<p>La respuesta correcta es la opción c), como muestra la siguiente deducción:</p> <p>1. Si el cliente de Teleplán puede darse de alta telefónicamente, puede contratar los servicios sin firmar ningún contrato.</p>

<p>periódicamente cada seis meses...</p> <p>Teleplán pone a su disposición para cualquier consulta el teléfono de atención gratuito 2002, la línea de atención comercial y la página web de teleplan.com.</p> <p><i>Si usted se ha dado de alta con nosotros telefónicamente, debe saber que tiene derecho a disponer de un contrato escrito en el que figuren las condiciones que se le aplican. Si no lo ha recibido, puede solicitarlo en teleplan.com o en el teléfono gratuito 900 456 667.</i></p>	<p>semestral de los cambios en las condiciones.</p> <p>c) puede contratar sus servicios sin haber firmado ningún contrato.</p>	<p>2. El cliente de Teleplán puede darse de alta telefónicamente.</p> <p>C. El cliente de Teleplán puede contratar sus servicios sin firmar ningún contrato.</p>
--	--	--

Nota: La respuesta correcta la conocemos porque se ofrece en las “Claves” del examen disponibles en el sitio web, pero el comentario sobre el tipo de inferencia solicitada es nuestro. Omitimos un ejemplo del nivel C2 porque no encontramos una tarea de opción múltiple en la prueba administrada el 19 de noviembre de 2011. Asimismo, el énfasis en los textos de la segunda columna es nuestro y lo usamos para realzar en qué fragmento se rastrea la deducción.

Fuente: Instituto Cervantes (2017).

A partir del análisis anterior podemos notar que los reactivos parecen solicitar un tipo de inferencia deductiva. Ejercicios como “Pierre escribe un correo electrónico sobre...” o “La cena será...”, en conjunto con la información que aparece en la columna dos muestran que estamos frente a deducciones, porque la información está completa y la conclusión, que en este caso es la respuesta correcta, no agrega nada nuevo a lo que ya está expresado en las premisas, es decir, la información de la columna dos. Y aunque no estamos afirmando que todas las pruebas de comprensión lectora requieren inferencias de corte deductivo, es interesante apreciar que este tipo de ejercicios son ejemplares, lo cual nos abre el camino para avanzar una propuesta que presentamos a continuación.

## 5. Inferencia extendida y racionalidad acotada

Es comprensible que en las pruebas de comprensión lectora se soliciten únicamente inferencias deductivas porque la deducción es una inferencia clásica que además es viable en términos de diseño. Sin embargo, esta no es una suposición inocente puesto que, como hemos visto en la sección anterior, la deducción no es la única inferencia que tenemos disponible y el enfoque clásico implica ciertos compromisos sobre nuestros criterios de racionalidad.

En efecto, las inferencias deductivas suponen exigencias muy altas. Por ejemplo, como argumenta Morado (2004), las inferencias deductivas parecen implicar que para que una persona sea racional, ésta debería tener tres características: 1) inferir fuera de contexto, como si las inferencias ocurrieran fuera del tiempo y el espacio; 2) tener recursos ilimitados y ser lógicamente omnisciente, como si la persona racional nunca admitiera contradicciones con las consecuencias lógicas de sus creencias; y 3) ser infalible y

consistente, como si nunca hubiera una necesidad de revisar las creencias pasadas o como si tuviéramos la garantía de que no habrá contradicciones si partimos de premisas consistentes. No obstante, como argumenta Morado, este ideal no corresponde con la realidad, lo cual sugiere que deberíamos adoptar una visión más generosa de inferencia: “Necesitamos un nuevo enfoque que considere de manera rigurosa, y si es posible formal, la estructura de la racionalidad en situaciones concretas, limitadas y falibles” (Morado, 2004, p. 318).

Todd y Gigerenzer (2000), por ejemplo, asumen una perspectiva descriptiva del comportamiento humano y argumentan que para entender la toma de decisiones y el razonamiento en el mundo real (es decir, el mundo de las personas que no tienen poderes sobrenaturales para razonar, ni conocimientos ilimitados, ni tiempo interminable (cf. Fonseca Patrón, 2016) es necesario comprender la noción de racionalidad acotada (*bounded rationality*), como fue entendida por Herbert Simon, y que puede ser caracterizada por las limitaciones de la mente humana y la estructura de los entornos en los que ésta opera. Dada esta noción de racionalidad acotada, Todd y Gigerenzer (2000) emplean el término de “racionalidad ecológica” para hablar de heurísticas ecológicas: una estrategia de resolución de problemas es ecológicamente racional en la medida en que se adapta a la estructura de un entorno (Todd y Gigerenzer, 2000).

Pero si las normas de racionalidad acotada y la necesidad de una inferencia extendida parecen razonables y, como hemos visto en el apartado anterior, parece que es normal que las pruebas de comprensión lectora utilicen inferencias de tipo deductivo, no resulta incómodo sugerir que el desempeño de los estudiantes en las pruebas de comprensión lectora inferencial podría verse afectado, porque los estudiantes son agentes con racionalidad acotada que llevan a cabo inferencias contextuales. Volveremos a esta propuesta después de mostrar los resultados experimentales.

## 6. Metodología

### 6.1. Muestra

Los participantes fueron  $n=187$  universitarios (102 mujeres y 85 hombres), pertenecientes a tres universidades distintas. La edad promedio fue de 19.5 años ( $SD \pm 1,5$ ). Al momento de la aplicación del instrumento, todos los participantes eran estudiantes de licenciaturas y semestres diversos, inscritos en cuatro áreas de formación académica, a saber: de Ciencias Económico-Administrativas: 54; de Ingeniería y Ciencias Exactas: 40; de Ciencias Sociales y Humanidades: 39; y de Ciencias Naturales y de la Salud: 54.

### 6.2. Instrumento

Nuestro instrumento es una prueba de comprensión lectora de opción múltiple compuesta por 12 reactivos correspondientes al producto de tres tipos de inferencia (deducción, inducción, abducción), dos tipos de longitud (corto, largo) y dos tipos de complejidad (simple, complejo). Cada reactivo tenía una pregunta asociada a cuatro opciones de respuesta (una sola correcta). A continuación se presentan tres reactivos de la prueba a manera de ejemplo (cuadro 2).

Cuadro 2. Ejemplos de reactivos del instrumento

NOTAS SOBRE EL REACTIVO	CUERPO DEL TEXTO	PREGUNTA Y OPCIONES DE RESPUESTA	JUSTIFICACIÓN DE LA RESPUESTA CORRECTA
<p><b>Reactivo 1</b></p> <p>Referencia del texto: Martínez Rica, J. P. (1979). <i>Enciclopedia de la vida animal</i>. México: Bruguera Mexicana de ediciones, S.A. Tomo 13, p. 1907.</p> <p>Texto largo y simple.</p> <p>Inferencia solicitada: deductiva.</p>	<p>El pájaro bobo real (o pingüino) “no construye un nido, sino que guarda su único huevo sobre sus pies, protegiéndolo con un repliegue de la piel abdominal. Sin embargo, permanece en un pequeño territorio en vez de ir y venir con el huevo. La puesta se inicia en noviembre y dura hasta abril. Después de confiar el huevo al macho, la hembra se ausenta por dos semanas en busca de comida, y de este modo rehace las reservas alimenticias perdidas al formar el huevo. En adelante, ambos progenitores se turnarán en el cuidado del huevo o polluelo.”</p>	<p>De acuerdo al texto anterior, qué se concluye necesariamente:</p> <p>A) que el pájaro bobo vive y muere en un pequeño territorio. B) que la incubación del pájaro bobo real dura seis meses. C) que el macho del pájaro bobo real no come hasta abril. D) que algún pájaro bobo real es buen constructor de nidos.</p>	<p>Llamemos <math>Q</math> al siguiente conjunto de enunciados que resumen el cuerpo del texto:</p> <ol style="list-style-type: none"> <li>1. El pájaro bobo real no construye un nido.</li> <li>2. El pájaro bobo real permanece en un pequeño territorio.</li> <li>3. El pájaro bobo real inicia la puesta (del huevo) en noviembre y dura hasta abril.</li> <li>4. Después de confiar el huevo al macho, la hembra se ausenta por dos semanas en busca de comida; en adelante, ambos progenitores se turnarán en el cuidado del huevo.</li> </ol> <p>...</p> <p>n. El pájaro bobo real guarda su único huevo sobre sus pies. n+1. El pájaro bobo real no va y viene con el huevo.</p> <p>a) Si la opción <math>A</math> es correcta, <math>A</math> debe ser consecuencia deductiva de <math>Q</math>; sin embargo, la opción <math>A</math> proporciona información que no pertenece a <math>Q</math> (la información relacionada con la vida y muerte). Por esto la opción <math>A</math> no puede ser la respuesta correcta.</p> <p>b) Si la opción <math>B</math> es incorrecta, <math>B</math> no debe ser consecuencia deductiva de <math>Q</math>; sin embargo, de 3 se sigue que, si el pájaro bobo real inicia la puesta (del huevo) en noviembre y dura hasta abril, entonces la</p>

incubación del pájaro bobo real dura seis meses, que es justamente una paráfrasis de B. Por esto la opción B es la respuesta correcta.

c) La opción C no es la respuesta correcta, como en el inciso a).

d) La opción D no es la respuesta correcta, como en el inciso a).

**Reactivo 2**

Referencia del texto:  
 Sampedro, J. (2002). Hombre y ratón comparten el 99% de sus genes. En *El País*. (Adaptación). Recuperado de [http://elpais.com/diario/2002/12/05/sociedad/1039042801\\_850215.html](http://elpais.com/diario/2002/12/05/sociedad/1039042801_850215.html)

Texto corto y simple.

Inferencia solicitada: no-deductiva (inductiva).

Una de las grandes sorpresas que deparó el análisis del genoma humano en 2001 fue que nuestra especie tenía unos 30,000 genes. Pero aún quedaba lo mejor. El genoma del ratón revela que este modesto roedor tiene también 30,000 genes, y que comparte al menos el 99% de ellos con el ser humano.

De acuerdo al texto anterior, qué es lo más probable que hagan los científicos en el futuro:

- A) estudiarán las diferencias entre los seres humanos y los ratones.
- B) invertirán en la crianza de ratones de laboratorio de forma masiva.
- C) podrán estudiar casi todas las enfermedades humanas en el ratón.
- D) se especializarán en las características fisiológicas de los ratones.

Para justificar la corrección de este reactivo, les solicitamos a cinco lectores expertos que lo resolvieran, ellos por probabilidad subjetiva eligieron como respuesta más frecuente la opción C.

Esta elección significaría que la conclusión más probable a partir del cuerpo del texto (en el que se enuncia la gran similitud genética entre los seres humanos y los ratones) es la opción C. De hecho, al conocerse tal similitud, parece altamente probable que se realicen distintos experimentos en el ratón que para el caso de los seres humanos no sería éticamente permitido; de esta forma, lo más probable que harían los científicos en el futuro sería que podrían estudiar casi todas las enfermedades humanas en el ratón.

Si bien las opciones A, B y D son probables, no representan la opción más probable en términos de lo que se sabe en la actualidad

sobre este tipo de experimentos.

<p><b>Reactivo 3</b></p> <p>Referencia del texto: Cortázar, J. (1988). Instrucciones para subir una escalera. En <i>Historias de cronopios y de famas</i>. Madrid: Alfaguara.</p> <p>Texto largo y complejo.</p> <p>Inferencia solicitada: no-deductiva (abductiva).</p>	<p>“... se comienza por levantar esa parte del cuerpo situada a la derecha abajo, envuelta casi siempre en cuero o gamuza, y que salvo excepciones cabe exactamente en el escalón. Puesta en el primer peldaño dicha parte, que para abreviar llamaremos pie, se recoge la parte equivalente de la izquierda (también llamada pie, pero que no ha de confundirse con el pie antes citado), y llevándola a la altura del pie, se le hace seguir hasta colocarla en el segundo peldaño, con lo cual en éste descansará el pie, y en el primero descansará el pie. (Los primeros peldaños son siempre los más difíciles, hasta adquirir la coordinación necesaria. La coincidencia de nombre entre el pie y el pie hace difícil la explicación. Cúidese especialmente de no levantar al mismo tiempo el pie y el pie).”</p>	<p>De acuerdo al texto anterior, qué actividad es la que mejor se explicaría con esas instrucciones:</p> <p>A) pedalear una bici. B) saltar una cuerda. C) escalar una cima. D) subir una escalera.</p>	<p>Para justificar la corrección de este reactivo, les solicitamos también a cinco lectores expertos que lo resolvieran, ellos por probabilidad subjetiva eligieron como respuesta más frecuente la opción D. Esta elección significaría que en el caso de este reactivo la opción que supone la hipótesis más probable a partir del cuerpo del texto en el que se explica detalladamente cómo ir subiendo peldaños con los pies (a manera de instrucciones) corresponde a la opción D puesto que se sabe que los peldaños son partes de una escalera.</p> <p>De esta forma se descartan como incorrectas las opciones A, B y C.</p>
--	--	---	--

Fuente: Elaboración propia.

### 6.3. Procedimiento

Se les solicitó a los estudiantes que firmaran una carta de consentimiento informado donde se les explicaba la confidencialidad con la que se conservarían sus datos personales. Posteriormente, se les leyeron las instrucciones de la prueba de forma grupal. Durante la prueba se les midió el tiempo, de tal manera que los estudiantes no excedieran 40 minutos para responderla. Las condiciones para resolver la prueba fueron favorables en relación al horario matutino y al lugar (salones de clase).

### 6.4. Análisis estadístico

El análisis se realizó con el paquete estadístico R versión 2.8.1 (R Development Core Team, 2008). Para establecer si la probabilidad de contestar correctamente a un reactivo estuvo afectada por la longitud, la complejidad y el tipo de inferencia utilizamos un modelo

lineal generalizado mixto (GLMM) con una distribución binomial y una función de enlace logit; en él incluimos, como variable dependiente, la probabilidad de responder correctamente; y como variables independientes: la longitud (factor con dos niveles: corto y largo), la complejidad (factor con dos niveles: simple y complejo) y el tipo de inferencia (factor con tres niveles: deductivo, inductivo y abductivo). Además, el modelo incluyó todas las interacciones entre estas tres variables. Consideramos, también, la identidad del participante como variable aleatoria debido a que los datos no son independientes entre sí y realizamos selección de modelos mediante remoción secuencial de términos que no tuvieran una significancia estadística.

## 7. Resultados

Encontramos un efecto de la longitud, la complejidad y el tipo de inferencia sobre la probabilidad de tener la respuesta correcta. En el cuadro 3 presentamos los estadísticos que permiten apreciar con claridad los valores de significancia entre las variables de interés.

Cuadro 3. Efecto de la longitud, la complejidad y el tipo de inferencia sobre la probabilidad de contestar correctamente

TÉRMINOS	COEFICIENTE	± ERROR ESTÁNDAR	$\chi^2$	P
Inferencia*longitud*complejidad	-1,935	0,226	79,004	<0,001
Inferencia*longitud	-0,329	0,105	9,944	0,002
Inferencia*complejidad	-0,534	0,104	26,842	<0,001
Longitud	0,817	0,053	71,844	<0,001
Complejidad	0,252	0,097	6,790	0,009
Inferencia	0,941	0,051	403,360	<0,001

Nota: Los asteriscos indican la interacción de los términos (variables) considerados en cada análisis. Las filas muestran los términos que no fueron removidos (remoción secuencial) por su significancia estadística. La única interacción rechazada fue: Longitud\*complejidad.

Fuente: Elaboración propia.

Tomando la complejidad del texto como la perspectiva desde la cual describiremos los resultados, observamos, por un lado, que en las preguntas que solicitaban una inferencia deductiva, cuando el texto era complejo y corto, 63 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,34; en contraste, con un texto complejo y largo, únicamente 40 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,21. Ahora bien, en las preguntas que requerían una inferencia abductiva, cuando el texto era complejo y largo, 124 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,66; en contraste, con un texto complejo y corto, únicamente 38 estudiantes contestaron correctamente, es decir, hay una probabilidad de 0,2. Por último, en las preguntas que promovían una inferencia inductiva, cuando el texto era complejo y largo, 176 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,94; en contraste, con un texto complejo y corto, 157 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,84. En suma, con textos complejos y largos, los estudiantes tienen más respuestas correctas si el texto solicita una inferencia no-deductiva.

Observamos, por otro lado, que en las preguntas que demandaban una inferencia deductiva, cuando el texto era simple y largo, 135 estudiantes contestaron correctamente, lo que significa una probabilidad de 0,72; en contraste, con un texto simple y corto, únicamente 29 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,16. Ahora bien, en las preguntas que solicitaban una inferencia abductiva, cuando el texto era simple y corto, 96 estudiantes contestaron de manera correcta, con una probabilidad de 0,51; en contraste, con un texto simple y largo, 45 estudiantes contestaron correctamente, lo que equivale a una probabilidad de 0,46. Por último, en las preguntas que pedían una inferencia inductiva, cuando el texto era simple y largo, 154 estudiantes contestaron correctamente, lo cual es una probabilidad de 0,82 que coincide con los resultados de cuando el texto era simple y corto, donde 153 estudiantes contestaron correctamente. En suma, con textos simples y largos, los estudiantes tienen más respuestas correctas si el texto solicita una inferencia inductiva, pero con textos simples y cortos, los estudiantes tienen menos respuestas correctas si el texto solicita una inferencia deductiva. En la figura 1 presentamos visualmente los resultados previamente descritos.

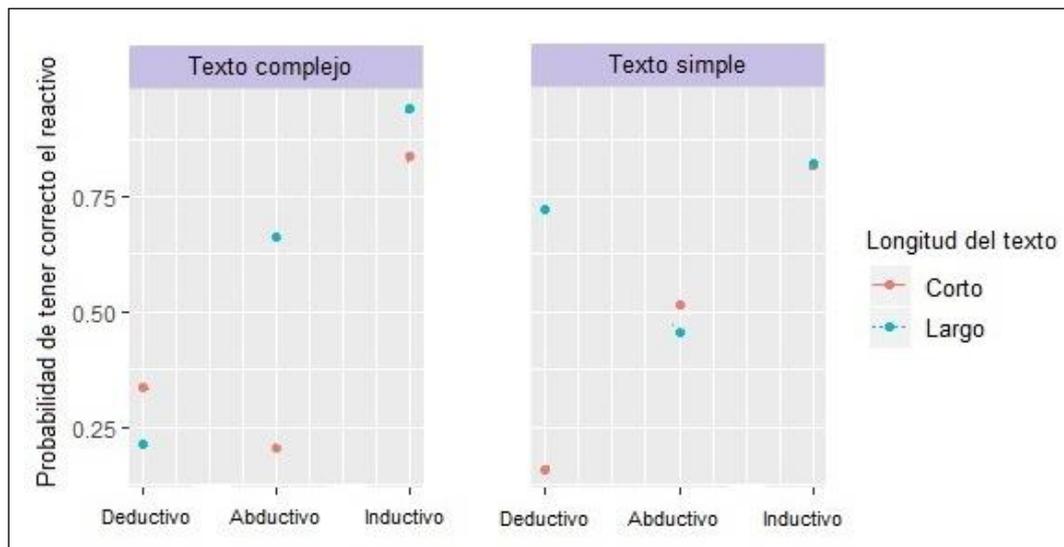


Figura 1. Efecto de la longitud, complejidad y tipo de inferencia sobre la probabilidad de responder correctamente.

Fuente: Elaboración propia.

## 8. Discusión y conclusiones

### 8.1 Discusión

Como se ha referido anteriormente, la comprensión lectora es un proceso que no se puede observar directamente (cf. Pérez Zorrilla, 2005; Fletcher, 2006) y por esto, para evaluarla, se suelen realizar pruebas para medir el desempeño de los estudiantes. En este trabajo evaluamos la comprensión lectora inferencial considerando los atributos estructurales de los textos y las propiedades estructurales de las pruebas. De esta forma, asociamos la longitud con la microestructura; la complejidad con la macroestructura; y los tipos de inferencia con la superestructura.

Con respecto a la longitud (microestructura), notamos que cuando un texto es corto (además de simple y deductivo) sólo el 16% de los estudiantes responde correctamente, lo cual podría explicarse por la alta *densidad* de información que posee un texto corto, lo que hace que el texto sea difícil de comprender; al respecto, Pallotti (2013) afirma que una idea errónea que se tiene comúnmente es la de pensar que para que un texto sea más simple baste con acortarlo, cuando en realidad, si se mantiene constante la cantidad de información, un texto más corto será más denso en términos de la información que presenta (en comparación con un texto largo) y la densidad conceptual es, por ello, una de las principales causas de dificultad.

Con respecto a la complejidad (macroestructura) observamos que cuando un texto es complejo (además de largo e inductivo) el 94% de los estudiantes lo responde correctamente, lo cual podría explicarse de la siguiente manera: cuando un estudiante se enfrenta a un texto complejo éste debe comprenderlo mediante su cohesión y coherencia (Ferrari, 2003), lo cual podría llevarlo a realizar relecturas más atentas en contraste con textos simples que lo llevarían a realizar lecturas menos atentas.

Con respecto a la inferencia (superestructura) notamos dos datos particularmente interesantes: 1) cuando un texto promueve una inferencia deductiva (además de simple y largo), el 72% de los estudiantes lo responde correctamente, lo que permite argumentar a favor de la presencia continua de textos deductivos en pruebas de comprensión lectora, lo cual es consistente con los diseños usuales como hemos visto, por ejemplo, a partir del análisis de la Tabla 1; pero también, 2) cuando un texto promueve una inferencia no-deductiva, específicamente inductiva, (además de ser complejo y corto), el 84% de los estudiantes lo responde correctamente, lo que nos permite argumentar a favor de la inclusión de textos no-deductivos (inductivos) en pruebas de comprensión lectora. Esto es consistente con los conceptos de racionalidad acotada e inferencia extendida: en efecto, el desempeño de los estudiantes se ve favorecido cuando incluimos un criterio más generoso de inferencia, lo cual está en sintonía con un concepto de racionalidad acotada.

## **8.2 Conclusiones**

Es razonable concluir que las características de microestructura (longitud), macroestructura (complejidad) y superestructura (inferencia) de una prueba de comprensión lectora inferencial influyen en el desempeño de los estudiantes, por lo cual resulta plausible reconsiderar el uso de inferencias no-deductivas en el diseño de pruebas de comprensión lectora, pues esto podría fortalecer los instrumentos de evaluación que utilizamos para medir el nivel de comprensión lectora de nuestros estudiantes.

Además, sabiendo que estas pruebas pretenden medir el grado de interacción entre el lector y el texto, podemos mencionar que las características del texto que hemos estudiado (longitud, complejidad e inferencia) están relacionadas con características del lector que hemos dado por supuestas. Por ejemplo, cuando un estudiante se encuentra inscrito en un programa de licenciatura en una universidad, posee un conjunto de conocimientos aprendidos durante los cursos en los niveles escolares inferiores. Estos conocimientos previos al momento de resolver una prueba de comprensión lectora inferencial ofrecen elementos a los estudiantes para obtener conclusiones a partir de enunciados en los cuales estas conclusiones no se encuentran referidas explícitamente. Incluso, un estudio futuro podría comparar el desempeño de los estudiantes en relación a su área de estudio de procedencia.

Por último, es necesario comentar que, debido a restricciones temporales, una limitación de este estudio fue el uso de 12 reactivos únicos para agotar todas las combinaciones posibles de las variables de interés sin repetición, por lo que, en trabajos futuros, podríamos estudiar el desempeño de los estudiantes con más reactivos para garantizar no sólo todas las combinaciones sino también repetición. Además, valdría la pena llevar a cabo una investigación similar a la que hemos presentado aquí, pero considerando, además de las propiedades estructurales de la prueba, las características de los estudiantes en términos de sus estrategias de lectura o de resolución de problemas (cf. Platas-García, Castro-Manzano y Reyes-Meza, 2016). Asimismo, parece muy conveniente que en estudios futuros se comparara el desempeño de los estudiantes en dos tipos de pruebas, a saber: una con preguntas únicamente deductivas y otra con preguntas deductivas y no-deductivas.

**Agradecimientos.** Nos gustaría agradecer a los evaluadores por sus precisas correcciones y útiles comentarios. Este trabajo fue parcialmente financiado por el proyecto UPAEP 30108-1008.

## Referencias

- Acosta, I. (2009). *La comprensión lectora, enfoques y estrategias utilizadas durante el proceso de aprendizaje del idioma español como segunda lengua*. Tesis doctoral. Universidad de Granada, España.
- Aliseda, A. (1997). *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. Amsterdam: Institute for Logic, Language and Computation, Universidad de Amsterdam.
- Aliseda, A. (1998). La abducción como cambio epistémico: C. S. Peirce y las teorías epistémicas en inteligencia artificial. *Analogía*, 12, 125-144.
- Anderson, R. C. y Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. En P. Carrell, J. Devine, D. Eskey. (Eds.), *Interactive approaches to second language reading* (pp. 37-55). Cambridge: Cambridge University Press.
- Bell, T. (2001). Extensive reading: speed and comprehension. *The Reading Matrix*, 1(1). Recuperado de <http://www.readingmatrix.com/articles/bell/>
- Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Estrasburgo: Consejo de Europa, Ministerio de Educación y Grupo Anaya.
- Cook, R. T. (2009). *A Dictionary of Philosophical Logic*. Edimburgo: Edinburgh University Press.
- Copi, I. y Cohen, C. (2007). *Introducción a la Lógica*. Ciudad de México: Limusa.
- Ferrari, S. (2003). Comprensione e lettura in L2: uno studio sperimentale sugli effetti della modificazione testuale. En R. Grassi, A. Valentini y R. Bozzone Costa. (Eds.), *L'italiano per lo studio nella scuola plurilingue: tra semplificazione e facilitazione* (pp. 85-110). Perugia: Guerra.
- Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading*, 10, 323-330.
- Fonseca Patrón, A. L. (2016). El debate sobre las heurísticas. Una disputa sobre los criterios de buen razonamiento entre la Tradición de Heurística y Sesgo y la Racionalidad Ecológica. *Valenciana*, 17, 87-115.

- García, E. (1993). La comprensión de textos. Modelo de procesamiento y estrategias de mejora. *Didáctica (Lengua y literatura)*, 5, 87-114.
- Giovannini, M. L. y Rosa, A. (2015). *Prove standardizzate di comprensione dei testi per la scuola secondaria di primo grado II. In uscita dalla classe seconda e in entrata nella classe terza*. Milano: LED.
- Guerra, J. y Guevara, Y. (2013). Validación de un instrumento para medir comprensión lectora en alumnos universitarios mexicanos. *Enseñanza e Investigación en Psicología*, 18(2), 277-291.
- Guerra, J. y Guevara, C. Y. (2017). Variables académicas, comprensión lectora, estrategias y motivación en estudiantes universitarios. *Revista Electrónica de Investigación Educativa*, 19(2), 78-90.
- Hawthorne, J. (2014). *Inductive Logic*. Recuperado de <http://plato.stanford.edu/archives/win2014/entries/logic-inductive>
- Instituto Cervantes. (2017). *Exámenes DELE. Preparar la prueba DELE*. Recuperado de <https://exámenes.cervantes.es/es/dele/preparar-prueba>
- Intraversato, A. (2013). *La comprensione della lettura fra abilità e conoscenze enciclopediche. Indagine sui Licei*. Roma: Ediciones Nuova Cultura.
- Mangen, A., Walgermo, B. R. y Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61-68. <https://doi.org/10.1016/j.ijer.2012.12.002>
- Morado, R. (2004). Problemas filosóficos de la lógica no monotónica. En R. Orayen y A. Moretti (Eds.), *Filosofía de la lógica, Enciclopedia Iberoamericana de Filosofía* (pp. 313-344). Madrid: Trotta y Consejo Superior de Investigaciones Científicas.
- Okasha, S. (2002). *Philosophy of Science: A very Short Introduction*. Oxford: Oxford University Press.
- Pallotti, G. (2013). Aiutare a capire e ad apprendere. En F. Pagnozzi y P. Weber (Eds.), *Testi narrativi per le classi con competenze linguistiche diversificate. Italiano come lingua seconda* (pp. 3-55). Trento: IPRASE.
- Pérez Zorrilla, M. J. (2005). Evaluación de la comprensión lectora: dificultades y limitaciones. *Revista de Educación, Ext.*, 121-138.
- Platas-García, A., Castro-Manzano, J. M. y Reyes-Meza, V. (2016). Razonamiento y heurísticas en pruebas de comprensión lectora. *Revista de Educación*, 371(1), 157-187. <https://doi.org/10.4438/1988-592X-RE-2015-371-312>
- R Development Core Team. (2008). *A Language and Environment for Statistical Computing*. Viena: R Foundation for Statistical Computing.
- Rice, G. E. (1994). Examining constructs in reading comprehension using two presentation modes: Paper vs. computer. *Journal of Educational Computing Research*, 11(2), 153-178. <https://doi.org/10.2190/MV46-VW49-4G5W-FM92>
- Rodríguez Rodríguez, R. (2005). Abducción en el contexto del descubrimiento científico. *Revista de Filosofía de la Universidad de Costa Rica*, 43(109/110), 87-97.
- Salazar, L. (2006). Interdependencia lingüística, transferencia y enseñanza-aprendizaje de lenguas extranjeras. *Laurus*, 12, 45-72.
- Sánchez Lobato, J. (Coord.). (2007). *Saber escribir*. Madrid: Aguilar/Instituto Cervantes.

- Serrano, M. A., Vidal-Abarca, E. y Ferrer, A. (2017). Decisiones estratégicas de lectura y rendimiento en tareas de competencia lectora similares a PISA. *Educación XXI*, 20(2), 279-297. <https://doi.org/10.5944/educXX1.12076>
- Solé, I. (1987). Las posibilidades de un modelo teórico para la enseñanza de la comprensión lectora. *Infancia y Aprendizaje*, 39-40, 1-13.
- Todd, P. M. y Gigerenzer, G. (2000). Précis of simple heuristics that make us smart. *Behavioral y Brain Sciences*, 23, 727-780.
- Van Dijk, T. A. (1992). *La ciencia del texto. Un enfoque interdisciplinario*. Barcelona: Paidós Comunicación.
- Zambelli, L. (2014). Semplificare i testi di studio: quando, come. *Italiano LinguaDue*, 1, 327-341.

## Breve CV de los autores

### Alejandra Platas-García

Licenciada en Filosofía y Maestra en Pedagogía. Actualmente estudia el Doctorado en Educación en la Universidad Popular Autónoma del Estado de Puebla (UPAEP). Es docente de italiano como lengua extranjera en el Centro de Lenguas Extranjeras de la Facultad de Lenguas de la Benemérita Universidad Autónoma de Puebla (BUAP). Realiza su investigación sobre la evaluación de la comprensión lectora inferencial a través de pruebas de opción múltiple. Asimismo, se interesa por las estrategias para la resolución de pruebas de comprensión lectora inferencial que emplean los estudiantes universitarios. ORCID ID: 0000-0002-5150-0668. Email: [aplatasg@gmail.com](mailto:aplatasg@gmail.com)

### J. Martín Castro-Manzano

Licenciado en Filosofía y Maestro en Inteligencia Artificial por la Universidad Veracruzana (UV), Doctor en Filosofía de la Ciencia por la Universidad Nacional Autónoma de México (UNAM); tiene estudios en Ciencias Computacionales por la BUAP. Actualmente es docente investigador de la Facultad de Filosofía y Humanidades en la UPAEP. Es miembro del Sistema Nacional de Investigadores (SNI, nivel 1), de la Academia Mexicana de Lógica y de la Association for Symbolic Logic. Realiza su investigación sobre Sistemas formales y Didáctica de la Lógica. ORCID ID: 0000-0003-2227-921X. Email: [josemartin.castro@upaep.mx](mailto:josemartin.castro@upaep.mx)

### Verónica Reyes-Meza

Licenciada en Psicología y Maestra en Diagnóstico y Rehabilitación Neuropsicológica. Doctora en Neuroetología, realizó un Posdoctorado en el posgrado en Ciencias Biológicas de la Universidad Autónoma de Tlaxcala (UATx). Actualmente es docente investigadora del Centro Tlaxcala de Biología de la Conducta en la UATx. Es miembro del Sistema Nacional de Investigadores (SNI, nivel 1) y de la Society for Neurosciences. Su principal línea de investigación es el análisis experimental de la conducta. ORCID ID: 0000-0002-2745-4032. Email: [veronica.reyesm@uatx.mx](mailto:veronica.reyesm@uatx.mx)

**Irene Gaona-Gordillo**

Licenciada en Biología por la Universidad Nacional Autónoma de México (UNAM) y Maestra en Ciencias Biológicas por la Universidad Autónoma de Tlaxcala (UATx). Es experta en análisis estadístico y programadora en paquetería R, específicamente en modelos generalizados con efectos mixtos. Realiza su investigación dentro de las áreas de Biología del Comportamiento y Ecofisiología. Email: r.gago.r@gmail.com

## Aplicación del Método Delphi Modificado para la Validación de un Cuestionario de Incorporación de las TIC en la Práctica Docente

### Application of the Modified Delphi Method for the Validation of a Questionnaire on the Incorporation of ICT in Teaching Practice

Carlos Enrique George Reyes \*<sup>1</sup>  
Laura Trujillo Liñan <sup>2</sup>

<sup>1</sup>Universidad Autónoma del Estado de Hidalgo - México

<sup>2</sup>Universidad Panamericana - México

La incorporación de las TIC en los escenarios educativos representa un reto para los docentes en el sentido de auto orientarse en la forma en la que utilizan estas herramientas en su práctica áulica e incluso en sus actividades cotidianas, cabe mencionar que estas formas de impactan directamente en la formación de los estudiantes, en la manera de conducirse a lo largo del proceso enseñanza-aprendizaje, así como en su formación personal. En este artículo se documenta el proceso de diseño y validación de un cuestionario para analizar cómo estos docentes están incorporando las TIC desde el referente teórico de la práctica docente. Para tal fin, se utilizó una variación del Método Delphi que permitió una respuesta eficiente de los expertos seleccionados para la validación. Como resultado se logró obtener un cuestionario con una escala tipo Likert en donde se refleja la experiencia profesional de expertos de distintas universidades para otorgar elaborar un cuestionario con claridad en las preguntas e ítems pertinentes.

**Palabras clave:** Práctica docente; TIC; Validez; Método delphi; Juicio de expertos.

The incorporation of ICT in educational scenes represents a challenge for teachers in a self-oriented way in which they use those tools in their classroom practice and even in their daily activities. It should be mentioned that these forms of use, make an impact on the formation of the students, in the way they conduct themselves throughout the teaching-learning process, as well as in their personal formation. The purpose of this paper is to document the design and validation process of a questionnaire to analyze how some teachers are incorporating ICT from the theoretical reference of a teaching practice. For this purpose, a variation of the Delphi Method was used, which allowed an efficient response of the experts selected for the validation. As a result, a questionnaire with a Likert-type scale was obtained and the professional experience of teachers from different universities to grant a questionnaire with clarity in the questions and pertinent items.

**Keywords:** Teaching practice; ICT; Validity; Delphi method; Experts judgment.

## 1. Introducción

Actualmente las instituciones educativas se encuentran inmersas en un cambio constante debido a la irrupción de las Tecnologías de la Información y la Comunicación (TIC) en el escenario del aprendizaje y la generación de conocimientos (Valencia et al., 2016), lo que ha originado que organizaciones como la OCDE (2006, 2016) y la UNESCO (2011) se pronuncien respecto a la importancia de la figura del docente como el responsable de incorporar estas herramientas con el objetivo de potenciar estos cambios e incluso mejorar la práctica docente. Lo anterior debido a que en la actividad áulica confluyen cuatro elementos muy importantes: el profesor, los estudiantes, el contenido y la tecnología que se usa para la enseñanza y el aprendizaje; la interacción que se da entre estos elementos, es lo que impactará en el aprendizaje.

Asimismo, esta irrupción también ha originado que las políticas educativas estén priorizando la dotación de infraestructura tecnológica desde hace muchos años (López y Ramírez, 2016), sobre la formación recursos humanos que sean capaces de lograr una integración pertinente de las TIC, esta situación si bien es cierto, beneficia a los docentes y a los alumnos respecto a las posibilidades para acceder a computadoras, software e internet, no garantiza que se estén desarrollando prácticas docentes innovadoras, sino que en la mayoría de los casos hace que se reproduzcan viejas concepciones pedagógicas con el uso de nuevos medios (Aparici, 2011).

Es por ello que se debe considerar al profesor como aquel que puede tomar ventaja de los beneficios de la tecnología en cada uno de sus cursos que imparte, ya que las TIC pueden ayudar a aumentar la eficiencia de estos, aumentar el compromiso y la satisfacción de los estudiantes y mejorar los resultados del aprendizaje cuando se usan de manera efectiva. Finalmente es el docente quien elige la tecnología que desea utilizar, ya sea para sustituir otra herramienta o para mejorar el acercamiento del alumno a ciertos conocimientos.

Es por lo anterior que en este artículo se propone la elaboración de un cuestionario que permita medir la incorporación de las TIC desde el enfoque de la práctica docente, es decir, desde la actividad que el docente desarrolla en el aula y que está delimitada por las distintas relaciones que se generan entre los docentes, los alumnos, la sociedad y la institución educativa de Fierro, Fortoul y Rosas (2000) y Hernández (2011), para ello, se diseñó el modelo de incorporación presentado en la figura 1.

A partir de este modelo se pueden categorizar los niveles de incorporación de las TIC de los profesores en relación con su práctica docente, teniendo como dimensiones contextuales la dimensión institucional y la dimensión social, ya que son indispensables para tener un marco de referencia normativo y del entorno en donde se desarrolla la actividad docente, de igual forma se agregan las dimensiones interpersonales, valorales, personales y didáctica, todas ellas unidas por la relación pedagógica-tecnológica que se manifiesta como la forma en la que el profesor vive la práctica docente a partir de su conocimiento de las dimensiones de su propia práctica y de las etapas por las que se encuentra transitando para incorporar las TIC a sus tareas cotidianas.

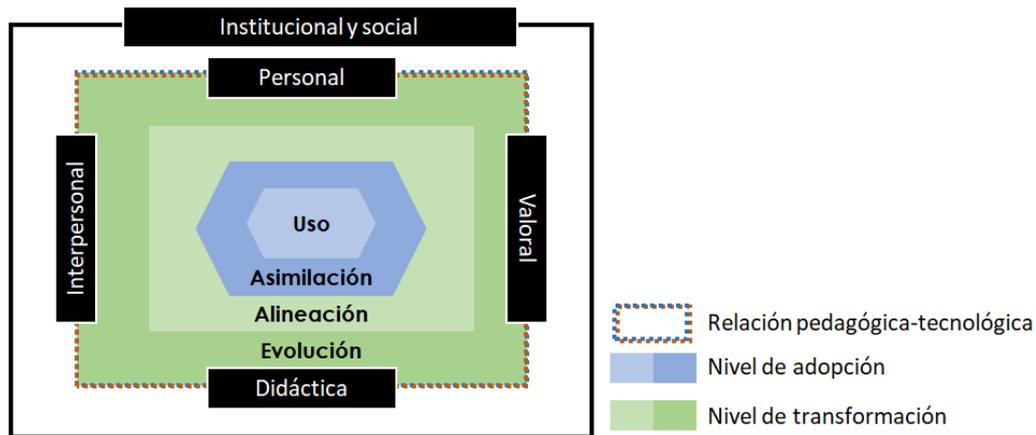


Figura 1. Modelo de incorporación de las TIC desde el enfoque de la práctica docente  
Fuente: Elaboración propia a partir de de Fierro, Fortoul y Rosas (2000), Hernández (2011) y Valencia et al. (2016).

## 2. Breve mirada del impacto antropológico del uso de las TIC en los docentes

La tecnología juega un papel muy importante en la educación hoy en día, como se mencionó anteriormente, los centros educativos buscan favorecer la educación de los alumnos adoptando nuevas herramientas tecnológicas para eficientar y favorecer la adquisición del conocimiento por parte de los estudiantes. Sin embargo se debe mencionar, que en esta adaptación es sumamente importante la participación del docente pues finalmente, él es el medio a través del cual los conocimientos llegan al alumno a través de las diferentes estrategias de enseñanza-aprendizaje que diseña.

Así, el incorporar las tecnologías al salón de clases no siempre es favorecedor debido a que el docente muchas veces no cree en la necesidad de incorporar nuevos recursos a la práctica docente, por falta de capacitación o simplemente por mantenerse en una zona de confort. Esto no es algo exclusivo del profesor, sino que en general, nos cuesta trabajo adaptarnos y adoptar nuevas formas de trabajo. Ya desde la antigua Grecia se hablaba de las nuevas tecnologías y su inserción en la sociedad:

Se cuenta que Theuth fue quien primero descubrió el número, el cálculo, geometría, astronomía, el juego de damas y sobre todo las letras. En aquella época, vivía el rey Thamus en la parte alta de Egipto. A él vino Theuth quien le mostraba sus artes y le decía que debían ser entregadas a todos los egipcios. Pero él preguntó, cuál era la utilidad que cada una tenía y mientras Theuth le decía los beneficios, Thamus aprobaba o desaprobaba su uso hasta que llegaron a las letras. Theuth afirmó al rey que ese conocimiento haría más sabios a los egipcios y más memoriosos, pues se ha inventado como un fármaco de la memoria y la sabiduría. A lo que Thamus respondió, que por amor a las letras, estaba atribuyendo poderes contrarios a los que tenía. Pues es olvido lo que producirán y descuidarán la memoria fiándose de lo escrito, llegarán al recuerdo desde fuera, a través de caracteres ajenos, no desde dentro, desde ellos mismos y por sí mismos (Fedro 274c-275a).

Con esto podemos ver que las nuevas herramientas pueden prometer eficiencia, rapidez e incluso mejorar nuestras actividades, en este caso en la enseñanza, pero la utilización de la misma también engloba una serie de consecuencias que debemos valorar en el momento de su implementación. Estos efectos tienen que ver con el impacto en el alumno, pues la información que llega a él en mucho depende del medio o tecnología que se utilice para su transmisión.

Un ejemplo de esto puede ser el uso de la memoria, a través de ella somos capaces de recordar hechos, pensamientos, ideas, que hemos vivido en el pasado. La memoria permite que haya en el hombre un pensamiento coherente y que se puedan realizar elevados procesos intelectuales, pues la mente es capaz de buscar en el archivo de la memoria y recuperar todo aquello guardado por la misma. Así, la memoria es un medio de la conciencia humana y de la mente (Trujillo, 2014).

En la antigua Grecia se creía que los alumnos debían memorizar textos completos pues éstos contenían los valores y costumbres de la cultura griega. Así, los estudiantes los memorizaban ayudados por el ritmo y el baile, de manera que, eran capaces de recitar textos de gran extensión para otras generaciones y de esta manera conservaban sus tradiciones de manera oral y reservadas solamente para aquellos que eran griegos (Havelock, 1963).

La educación de hoy en día tiene otras herramientas para auxiliar a los alumnos en el proceso de enseñanza-aprendizaje y ya no es oral, sino más bien escrita. Hoy se favorece que los alumnos lean y escriban, es un proceso más solitario, pues es necesario estar solo con el libro para poder absorber más de él. También, la escritura nos ha permitido ya no tener que memorizar normas, leyes, números, pues podemos acceder a esta información en cualquier momento.

Es así que los medios que se usan para el aprendizaje se han ido modificando a lo largo de la historia y así, las sociedades, las TIC en mucho favorecen el cambio en la manera de aprender y en la manera en que se transmite la información por ello, es decir, favorecen la transformación de la práctica docente, por lo que es importante atender qué medios o TIC deben usarse y para qué, pues esto repercutirá de manera esencial en el desarrollo de los educandos.

### **3. Diseño metodológico**

#### **3.1. Objetivos**

En esta investigación se documenta el proceso mediante el que se diseñó y validó un cuestionario capaz de analizar la incorporación de las TIC desde el enfoque de la práctica docente, los objetivos planteados para realizar lo anterior, son los siguientes:

1. Diseñar un cuestionario llamado “incorporación de las TIC en la práctica docente”.
2. Analizar la validez de contenido del cuestionario mediante la aplicación del Método Delphi.

### **3.2. Diseño del cuestionario y elaboración de ítems**

De acuerdo a D'Ancona (1999), el cuestionario es la aplicación de un procedimiento estandarizado para recabar información de un grupo de sujetos, que tiene como objetivo buscar información sobre la conducta, los valores y las actitudes respecto a un tema de investigación, el cuestionario se diseñó utilizando una escala tipo Likert, debido a que este tipo de escalas tiene la capacidad de medir la intensidad de una actitud, lo que permite ubicar a una persona en algún punto del continuo de una actitud (Kerlinger y Howard, 2002), asimismo, los cuestionarios realizados con este tipo de escala han mostrado tener un rendimiento muy aceptable en las experiencias investigativas (Blanco y Alvarado, 2005), en parte debido a que la puntuación por ítem del cuestionario permite precisar un mayor o menor grado de presencia de las variables de estudio en una investigación.

De igual forma, se ha demostrado que la percepción de los docentes hacia el uso de las TIC es un factor esencial para determinar cómo las utilizan (Valdés et. al., 2011), por lo que bajo esta premisa se elaboraron ítems a partir de una escala fijada estructuralmente por dos extremos recorriendo un continuo desde desfavorable hasta favorable para cada afirmación (Likert, 1932 citado por Casas, 1999).

Respecto a las opciones de respuesta de esta escala, Jay y Swerdlik (2006) mencionan que estas pueden adoptar muchas formas y se pueden hacer de varias maneras, afirmando que “ningún tipo de escala es la mejor” (p. 197) sino que deben diseñarse para adaptarse a lo que se desea medir, en este sentido, concluyeron que en general una escala de tipo Likert funciona mejor asignando valores de 1 a 5 para la aprobación de opciones en uno y otro extremo de la escala, sin embargo posteriores aplicaciones de la escala (Domínguez, et. al. 2013; Lozano, García-Cueto y Muñiz, 2008; Preston y Colman, 2000) han demostrado que el número óptimo de opciones de respuesta es de 4 a 7, mientras que escalas con menos de 4 y con más de 10 opciones disminuyen la confiabilidad del cuestionario.

Considerando lo anterior, se seleccionaron cuatro opciones de respuesta debido a que “si hay cuatro opciones en la escala (nunca, casi nunca, algunas veces y siempre) para cualquier enunciado, la distancia psicológica entre nunca y casi nunca es la misma que entre casi nunca y algunas veces” (Domínguez, 2013, p. 31), es decir, existe una simetría entre las respuestas que se pueden seleccionar, de igual forma si se elimina una respuesta intermedia para procurar que el respondiente incline su respuesta y con ello se pueden encontrar resultados más confiables.

Esta apreciación se refuerza empíricamente con la investigación realizada por Chang (1994), que demostró que una escala de 4 opciones tiene más fiabilidad que escalas de 5 opciones o más, de la misma forma (Lee y Paek, 2014), mencionan que al utilizar escalas de 4, 5 o 6 opciones se mantienen resultados poco significativos respecto a la fiabilidad de un instrumento, argumentando que “las propiedades psicométricas [en una escala] tienden a ser mejores con más puntos de respuesta pero, los beneficios de tener más categorías de respuesta tienden a estancarse en 4 puntos, por lo que parece que existe un intervalo óptimo (es decir, entre 4 y 6 puntos) en lugar de un punto de escala óptimo” (p. 669).

Por lo que, se decidió que las opciones del cuestionario contara con valores de 1 a 4, con una asignación de: 1 “nunca”, 2 “Algunas veces”, 3 “casi siempre” y 4 “siempre” y con una interpretación de la valoración de las opciones de acuerdo a las etapas de incorporación de

las TIC de en el desarrollo de la práctica docente de la acuerdo a la correspondencia 1 “uso”, 2 “asimilación”, 3 alineación y 4 “evolución” para medir las dimensiones de la práctica docente con el uso de las TIC.

#### **4. Validación del cuestionario con el método Delphi**

La validación o validez de contenido de un cuestionario es entendida como “el grado en el que se incluyen todos los elementos representativos de lo que se quiere medir” (Gil y Pascual, 2012, p. 3), al respecto se han sugerido distintas formas para evaluar si un cuestionario tiene una alta validez de contenido, sin embargo el método más habitual se basa en el trabajo con jueces expertos en el tema que se está investigando (Sireci, 1998), este método, llamado Delphi, ha sido utilizado para estructurar un proceso de comunicación grupal que consiste de una serie de preguntas que se envía a un conjunto de personas que debido a su experiencia profesional tienen la capacidad de emitir juicios de valor para que, como un todo puedan tratar problemas complejos (Silva et al., 2002).

Por lo que el objetivo de este método es conseguir un consenso fiable entre las opiniones de expertos en un tema de estudio específico, cabe destacar que las principales ventajas de este método se concentran en la factibilidad para que expertos de diferentes ramas de un tema de investigación pueden realizar la valoración, de igual forma se genera una interacción controlada de los participantes, además de que se mantiene el anonimato de los jueces, para que de esta forma se pueda recuperar una opinión que no esté mediada por la reputación o la influencia de un miembro del grupo (Rosas, Sánchez y Chávez, 2012),

Cabe mencionar que este método ha demostrado su eficacia siempre que ha sido utilizado (Christie y Barela, 2005), además de que ha sido utilizado de forma consistente en el terreno de la educación (Cabero e Infante, 2014; Cabero et al., 2008), así como en el campo de las TIC (Barroso y Cabero, 2011; Romero et al., 2012), por lo que es pertinente aplicarlo en esta investigación, de igual forma Cabero (2014) recomienda este método de validación cuando se desea mantener la heterogeneidad de los jueces a fin de asegurar la validez de los resultados, así como para hacer posible participar a expertos que se encuentran físicamente dispersos.

Como todo método, este sigue un procedimiento que de acuerdo con LinstoneyTuroff, (1975), Landeta (1999) y Aponte, Cardozo y Melo (2010), se puede resumir en ocho etapas: 1) la identificación del problema, 2) formación de un panel de expertos, 3) presentación del problema por medio de un cuestionario, solicitando respuesta, 4) las respuestas se sintetizan en una serie de afirmaciones, 5) las afirmaciones son enviadas al panel con el objetivo de considerar su primera opinión, 6) el panel responde y analiza la información para que el investigador, 7), en consecuencia, se identifican convergencias o divergencias para lograr un consenso respecto a las valoraciones obtenidas para, y 8) elaborar el cuestionario final. Gráficamente el procedimiento puede organizarse de la siguiente forma (figura 2).

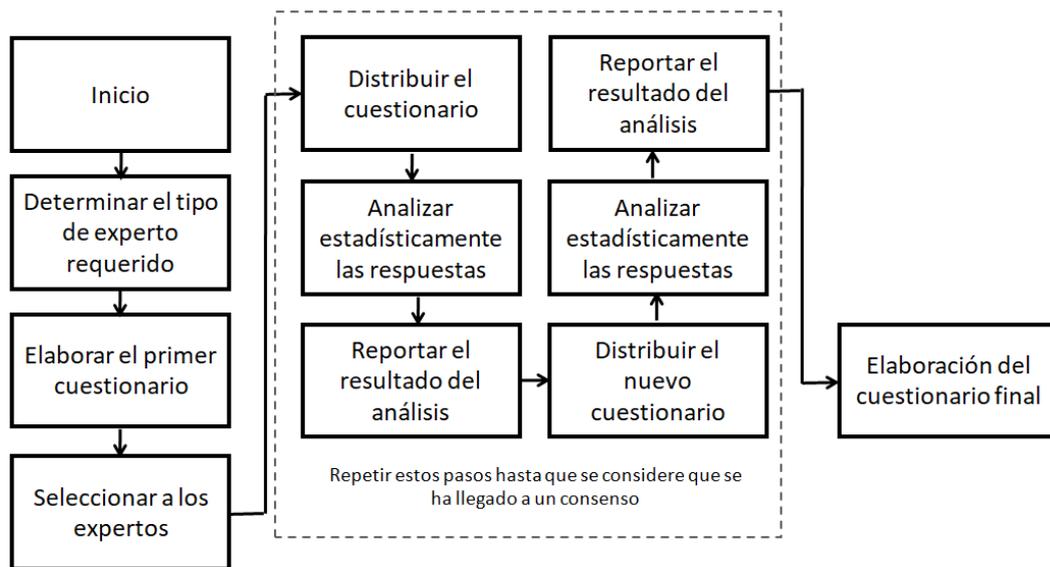


Figura 2. Proceso general del método Delphi

Fuente: Elaboración propia a partir de: Linstone y Turoff, (1975), Landeta (1999) y Aponte, Cardoso y Melo (2010).

#### 4.1. Selección de expertos

La selección de expertos es una de las etapas fundamentales en el método Delphi, ya que de la sección adecuada dependerá la confiabilidad de los resultados a obtener (Aponte, Cardoso y Melo, 2010), una buena selección podrá poner sobre la mesa opiniones argumentadas que no necesariamente serán unánimes, pero que permitirán identificar fortalezas y deficiencias en el instrumento que se esté evaluando, lo que dará lugar a tomar decisiones respecto a que ítems modificar o eliminar.

Fue importante definir el perfil de los expertos a considerar, para esta investigación se determinó que los miembros del jurado deben reunir al menos una de las siguientes características: 1) tener formación doctoral en Ciencias de la Educación o en áreasafines al tema de investigación, 2) ser experto en temas relacionados con las TIC y laEducación, con el grado de Maestría, de igual forma se buscó que los expertos provinieran de diversos contextos, con el fin de recuperar respuestas heterogéneas que sirvieran para obtener diferentes miradas acerca del cuestionario.

En cuanto al número de expertos, existen diversas opiniones que sitúan la cantidad entre 7 a 30 individuos (Landeta, 2002) a cantidades de 15 a 25 (García y Fernández, 2008), sin embargo, Cabero y Barroso (2013) mencionan que la selección de expertos en ocasiones no puede condicionarse a un rango predeterminado, debido a que muchas veces no se está en posibilidad de acceder a expertos suficientes con referencias claras a la temática analizada, en otras, porque se requiere rapidez para ofrecer resultados preliminares en una investigación o bien debido a que se quiera evitar la “mortandad experimental” (p. 28) por la desmotivación en la participación de los expertos en varias rondas de valoración.

Bajo estos criterios se estableció contacto con 27 expertos de diversas instituciones, de los cuales 16 se mostraron interesados por participar, pero solo 12 respondieron de forma favorable, como se puede apreciar en la siguiente tabla, el 91,67% corresponde a expertos

con grado académico de doctorado, los cuales cumplen con la primer característica deseada para la selección, mientras que el 8,33% corresponde a la característica relacionada de ser experto en el tema de las TIC y Educación y tener el grado de Maestría.

De igual forma, el 25% son expertos que laboran en universidades extranjeras y él 58,33% laboran en universidades mexicanas, mientras que un experto labora en una universidad privada de reconocido prestigio y otro experto realiza actividades exclusivamente directivas, asimismo el 83,33% son docentes de tiempo completo y dedicados a actividades investigativas, el siguiente cuadro muestra la distribución de los expertos.

Cuadro 1. Expertos que participaron en la validación por el método Delphi

NÚM.	INSTITUCIÓN A LA QUE PERTENECE	CATEGORÍA DE LA INSTITUCIÓN	GRADO ACADÉMICO	CATEGORÍA DOCENTE
S1	Instituto Nacional para la Evaluación de la Educación, Delegación Hidalgo, México	Organismo Público	Doctora en Ciencias de la Educación	Directora
S2	Universidad Autónoma del Estado de Hidalgo, México	Universidad Pública	Doctor en Psicología	Profesor Investigador
S3	Universidad Autónoma del Estado de Hidalgo, México	Universidad Pública	Doctor en Ciencias Técnicas	Profesor Investigador
S4	Universidad Autónoma del Estado de Hidalgo, México	Universidad Pública	Doctor en Geografía	Profesor Investigador
S5	Universidad Autónoma del Estado de Hidalgo, México	Universidad Pública	Doctora en Ciencias de la Computación	Profesor Investigador / Coordinadora
S6	Universidad de Cienfuegos, Cuba	Universidad Pública	Doctor en Ciencias Pedagógicas	Profesor Investigador
S7	Universidad Politécnica de Tulancingo	Universidad Pública	Doctora en Ciencias de la Educación	Profesor Investigador
S8	Universidad de Cienfuegos, Cuba	Privada	Doctora en Ciencias Pedagógicas	Profesor Investigador
S9	Universidad Nacional Autónoma de México (ISSUE), México.	Universidad Pública	Doctora en Educación Superior	Profesor Investigador
S10	Universidad Nacional Autónoma de México (ISSUE), México.	Universidad Pública	Maestría en Investigación Educativa	Doctorante en Investigación Educativa
S11	Universidad Abierta, Universidad Panamericana, México.	Privada	Doctora en Historia del Pensamiento	Profesor Investigador / Directora
S12	Universidad Tecnológica Nacional, Facultad Regional Mendoza, Argentina.	Universidad Pública	Doctor en Innovación Curricular, Tecnológica e Institucional	Profesor Titular

Fuente: Elaboración propia.

#### **4.2. Procedimiento de aplicación del Método Delphi**

En la aplicación del método Delphi original se realizan tres o más rondas de revisión, sin embargo se puede llevar a cabo una aplicación del Método Delphi en la cual se realizan dos rondas de participación con el objetivo de evitar que la validación se convierta en una tarea larga y costosa, así como para evitar la deserción de la respuesta de los expertos en rondas consecuentes (Cabero, 2014).

Por lo que la aplicación del método Delphi modificado se llevo a cabo en dos fases, en la primera se facilitó a los expertos una lista inicial de 50 ítems, obtenidos a partir de la revisión de la literatura correspondiente a las variables de investigación, práctica docente e incorporación de las TIC en la Educación Superior. El cuestionario quedo constituido por dos secciones, la primera referente a los datos generales que se desean recuperar el docente así como las instrucciones para responderlo y la segundo en donde se ubicaron los ítems a valorar.

Se utilizó una valoración de la representatividad de los ítems usando un formato de respuesta tipo Likert de 5 opciones, en donde la tarea de cada juez fue evaluar la relevancia de cada ítem del cuestionario, por lo que a cada uno se le proporcionó un cuestionario para que asignara un valor numérico a cada ítem, en donde el valor 0 (cero) representa “nada relevante” y 5 (cinco) representa “totalmente relevante”.

Posteriormente y de acuerdo a la sugerencia de Abad y otros (2011), se obtuvo la media y la desviación estándar para conocer el índice de relevancia de cada ítem, en donde si el ítem tiene un índice inferior a 0,75 es eliminado del cuestionario, los ítems que superan este índice son incluidos en el cuestionario final y son recolectados en un nuevo instrumento que se distribuye a los expertos para una segunda validación.

En la segunda fase y con el cuestionario rediseñado se procuró hacer más dinámico el método, abriendo la posibilidad de recuperar comentarios respecto a la calidad de la redacción de cada ítem, en consecuencia se invitó nuevamente a los jueces a emitir notas al margen del cuestionario para valorar otras posibles argumentaciones que pudieran enriquecer su confección para lograr una mayor validez de contenido, en resumen, el procedimiento para realizar el método Delphi adaptado consistió en: a) elaborar el cuestionario a partir de las variables de la investigación, b) seleccionar a los expertos, c) distribuir el cuestionario, d) analizar los resultados con base en medios estadísticos, e) rediseñar el cuestionario y reportar el resultado a los expertos, f) distribuirlo para recuperar notas al margen respecto a aspectos no identificados en la valoración estadística, g) analizar las notas recuperadas y h) elaboración definitiva del cuestionario, en proceso puede sintetizarse en la figura 3.

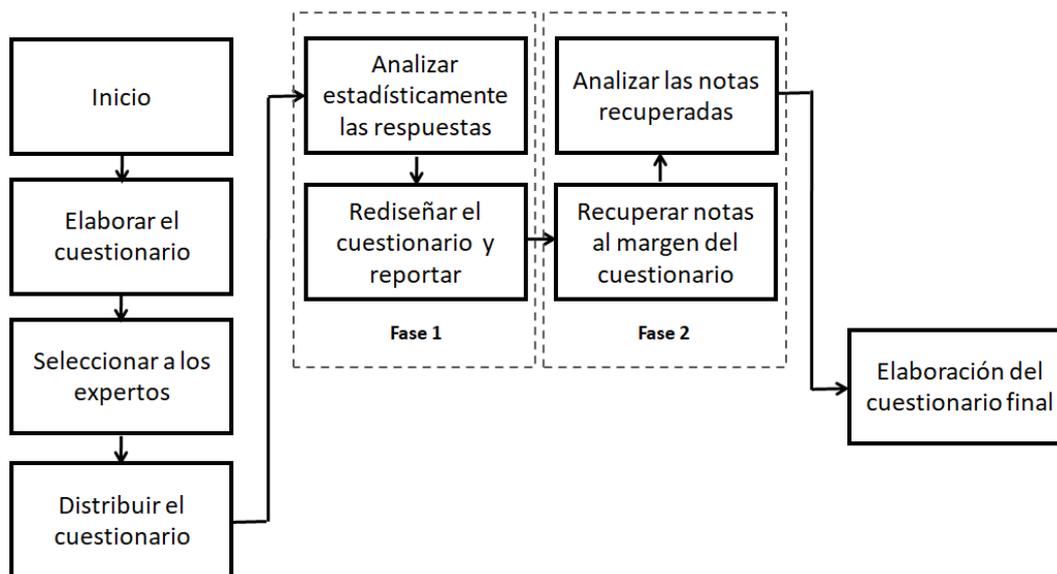


Figura 3. Propuesta de procedimiento de aplicación del método Delphi

Fuente: Elaboración propia.

## 5. Resultados

### 5.1. Primera fase

Una vez elaborado el cuestionario, seleccionado a los expertos y habiéndolo distribuido, se realizó la primera fase del método Delphi, que consistió en el análisis estadístico para obtener la media y la desviación estándar, a partir de la cual se tomaron decisiones respecto a mantener o eliminar ítems, se debe recordar que el criterio para mantener un ítem es que rebase el 75% de la media de acuerdo a las valoraciones de los jueces, que en una escala de 0 a 5 representa un índice de 3,75 puntos, en caso de que la relevancia sea menor a 3,75 puntos, se desecha, en la siguiente tabla se presentan los resultados para cada una de las dimensiones que constituyen el cuestionario (cuadro 2).

Cuadro 2. Ítems del cuestionario

ÍTEMS	MEDIA	DESVIACIÓN ESTÁNDAR	PERMANENCIA
<b>Dimensión Personal</b>			
A11) Uso herramientas o aplicaciones tecnológicas en mi vida cotidiana	2,583	1,256	No
A12) Mi formación profesional (licenciatura, maestría y/o doctorado ha incluido formación para utilizar las TIC	4,667	0,471	Si
A21) Participo en programas de formación relacionados con la utilización de las TIC en mi actividad docente, que son ajenos a la universidad (externos)	3,833	0,799	Si
A22) He participado en la formación de grupos de profesores que buscan utilizar las TIC en la docencia	1,917	0,862	No
A31) Comparto con otros docentes estrategias para incorporar las TIC en el aula	4,000	0,707	Si
A32) Estoy satisfecho por la forma en la que he incorporados las TIC en mi práctica docente	4,250	0,722	Si

ÍTEMS	MEDIA	DESVIACIÓN ESTÁNDAR	PERMANENCIA
<b>Dimensión institucional</b>			
B11) Conozco las tendencias de la UNESCO para incorporar las TIC en la Educación Superior	2,000	1,155	No
B12) Conozco el modelo educativo que regula la incorporación de las TIC en la universidad	3,750	1,534	Si
B21) Comparto la cultura de incorporación de las TIC en la universidad	4,333	0,850	Si
B22) La incorporación de herramientas tecnológicas en mi práctica docente corresponde con las lineamientos del modelo educativo de la universidad	4,333	0,745	Si
<b>Dimensión didáctica</b>			
C11) Uso de forma cotidiana dispositivos digitales como laptops, o tablets en la actividad docente	4,417	0,640	Si
C12) Uso herramientas o aplicaciones tecnológicas en mi vida cotidiana	3,250	1,479	No
C13) Identifico las características comunes de herramientas tecnológicas como menús de navegación opciones y configuraciones	3,833	1,143	Si
C21) Uso correo electrónico para comunicarme con mis compañeros y alumnos	4,500	0,500	Si
C22) Uso Skype, Hangouts u otra herramienta de comunicación inmediata para estar en contacto con mis compañeros y alumnos	4,500	0,500	Si
C31) Participo con mis alumnos en redes sociales virtuales como Edmodo, Educanetwork, LinkedIn	4,583	0,493	Si
C32) Participo en redes profesionales virtuales como REDOLAC, Eduteka, Red Docente OEI	4,667	0,471	Si
C41) Desarrollo de actividades educativas en plataforma educativa de la universidad (Blackboard o Moodle)	4,667	0,471	No
C42) Diseño de escenarios de comunicación virtuales como grupos en Gdrive, Google Classroom o portales educativos	4,250	1,090	Si
C51) Utilizo bases de datos y fuentes de información para realizar actividades de investigación en mi práctica docente	4,833	0,373	Si
C52) Busco y filtro información de internet usando las opciones de navegadores de internet como Chrome o Mozilla	4,750	0,595	Si
C61) Divulgo los resultados de mis actividades de investigación en escenarios como congresos virtuales o publicaciones en internet	4,500	0,645	Si
C62) Diseño escenarios de aprendizaje virtuales como blogs y páginas web	4,500	0,645	Si
C71) Participo en comunidades virtuales de formación como MOOCS, cursos en línea, seminarios virtuales	4,583	0,493	Si
C81) Diseño materiales digitales con herramientas tecnológicas interactivas como Ardora, Educaplay, eXeLearning, Prezi	4,667	0,471	Si
C82) Combino herramientas tecnológicas en la planificación de la actividad docente tales como presentaciones con organizadores gráficos, videos con documentos, herramientas colaborativas con gestores bibliográficos, etc.	4,750	0,433	Si

ÍTEMS	MEDIA	DESVIACIÓN ESTÁNDAR	PERMANENCIA
<b>Dimensión interpersonal</b>			
D11) Mis alumnos me demandan que incorpore las de las TIC en las actividades de aprendizaje	4,917	0,276	Si
D12) Atiendo las demandas de que mis alumnos hacer respecto a utilizar las TIC	4,750	0,433	Si
D13) Oriento a mis alumnos para incorporar las TIC en sus actividades de aprendizaje	4,917	0,276	Si
D21) Comparto con otros docentes estrategias para incorporan las TIC en el aula	4,583	0,640	Si
D22) Participo en proyectos de enseñanza relacionados con la utilización de las TIC	4,167	1,067	Si
<b>Dimensión social</b>			
E11) La universidad cuenta con la infraestructura necesaria para que incorpore las TIC en mi práctica docente	4,583	0,640	Si
E12) Cuento de forma personal con dispositivos y conectividad para incorporar las TIC en mi práctica docente	4,667	0,471	Si
E21) Uso de forma cotidiana dispositivos digitales como laptops, o tablets en la actividad docente	4,917	0,276	Si
E31) Me percibo como un docente que transforma su práctica al utilizar las TIC en la enseñanza	5,000	0,000	Si
<b>Dimensión Valoral</b>			
F11) La incorporación de las TIC en la educación está relacionada con la mejora en de la práctica docente	4,833	0,373	Si
F12) La incorporación de las TIC en la educación pueden mejorar las estrategias de enseñanza	4,417	0,862	Si
F21) Propongo a mis alumnos proyectos de aprendizaje en donde deben incorporan las herramientas tecnológicas	4,417	0,862	Si
F22) En el aula, me siento cómodo con el uso de las tecnologías en general	4,667	0,624	Si
F31) De acuerdo con mi experiencia, incorporar en la práctica educativa el uso de TIC podría mejorar los logros académicos	4,833	0,373	Si
F32) Identifico en mi práctica docente problemáticas educativas que los alumnos pueden resolver usando las herramientas tecnológicas	4,917	0,276	Si
F41) Aplico métodos de diseño instruccional apoyados en herramientas tecnológicas para planear mi práctica docente	3,417	1,847	Si
F51) La incorporación de las TIC en la práctica docente obedece a intereses políticos	5,000	0,000	Si
F52) La incorporación de las TIC en la práctica docente obedece a intereses académicos	5,000	0,000	Si
<b>Relación pedagógica-tecnológica</b>			
G11) La generación de conocimientos se enriquece al incorporar las TIC en la educación	4,750	0,433	Si
G21) Considero que mis alumnos podrían tener un buen nivel de satisfacción durante el proceso de enseñanza-aprendizaje si está mediado por tecnologías	4,667	0,471	Si
G31) De acuerdo con la información y conocimiento con el que cuento, la práctica docente podría verse favorecida con el uso de TIC	4,833	0,373	Si

ÍTEMS	MEDIA	DESVIACIÓN ESTÁNDAR	PERMANENCIA
G32) Me satisface incorporar estrategias educativas que incluyan aplicaciones tecnológicas	3,750	1,920	Si
G41) Comparto con otros docentes estrategias para incorporan las TIC en el aula	3,250	2,314	No
G42) Al elaborar estrategias pedagógicas incluyo a las TIC	3,583	1,847	No
G43) Incorporo herramientas tecnológicas en mi plan de clase de acuerdo a los objetivos de aprendizaje que requieren el uso de las TIC	4,917	0,276	Si

Fuente: Elaboración propia.

Después de analizar los resultados estadísticos y de acuerdo con el índice de relevancia ya mencionado, se decidió eliminar los ítems A11 (2,583) A22 (1,917), B11(2,000), C12 (3,250), F41 (3,417), G41 (3,250) y G42 (3,583) por no haber rebasado los 3,75 puntos necesarios para permanecer en el cuestionario, lo que permitió elaborar una nueva versión del cuestionario para proseguir con el método.

### 5.2. Segunda fase

En la segunda fase, se distribuyó el cuestionario a los expertos para solicitar una valoración cualitativa a través de notas al margen con el objetivo de conocer propuestas argumentativas para mejorar la calidad de la redacción de los ítems, haciendo énfasis en ítems que obtuvieron valores mayores a 1 punto en la medición de la desviación estándar, es decir, en ítems que reflejaron una mayor dispersión o variabilidad en su valoración por parte de los expertos, los cuales fueron B11 (1,155), B12 (1,534), C12 (1,479), C13 (1,143), C42 (1,090), D22 (1,067), G32 (1,920), G41 (2,314) y G42 (1,847).

Las valoraciones cualitativas realizadas dieron como resultado la reformulación de algunos ítems, ya fuera porque las notas mencionaban insuficiencia en la calidad de la redacción, ambigüedad o generalidad o bien porque argumentativamente tenían una propuesta que a su parecer pudiera dar cuenta de mejor manera lo que se deseaba medir, cabe mencionar que en esta fase participaron 5 expertos, en la siguiente tabla se muestran las valoraciones realizadas y la redacción final del ítem (cuadro 3).

Cuadro 3. Modificaciones y actualizaciones de ítems

ÍTEM ORIGINAL	ÍTEM MODIFICADO
A12) Mi formación profesional (licenciatura, maestría y/o doctorado ha incluido formación para utilizar las TIC	A12) Mi formación profesional (licenciatura, maestría y/o doctorado ha incluido formación para utilizar las TIC en mi actividad docente
A21) Participo en programas de formación relacionados con la utilización de las TIC en mi actividad docente, que son ajenos a la universidad (externos)	A21) Participar en programas de formación relacionados con la utilización de las TIC me sirven para desarrollar de mejor manera mi actividad docente.
<i>S1: "Participo" Si/No, evita este tipo de redacción en los ítems</i>	
<i>S2: ¿Qué relevancia tiene que sean externos? Sugiero modificar el ítem y quitar lo de externo</i>	
A31) Comparto con otros docentes estrategias para incorporar las TIC en el aula	A31) Colaboro con otros docentes compartiendo estrategias

ÍTEM ORIGINAL	ÍTEM MODIFICADO
<p>Notas:</p> <p><i>S1 ¿Cómo cuáles? Comparto con otros docentes ¿Qué estrategia? Te sugiero usar colaborar en vez de compartir</i></p> <p>A32) Estoy satisfecho por la forma en la que he incorporados las TIC en mi práctica docente</p> <p><i>S1: Mide satisfacción, pero no queda claro como mide la práctica docente, quita satisfacción, busca un sinónimo menos subjetivo</i></p> <p><i>S2: Esto es satisfacción, no sé si es suficiente para tu objetivo</i></p> <p><i>S3: Sugiero que la "satisfacción" se cambie por algún tipo de resultado", algo medible, no algo relativo</i></p>	<p>aprendidas para utilizar las TIC en el aula tales como definir objetivos para usar las TIC o usar entornos virtuales para el aprendizaje.</p> <p>A32) Utilizar las TIC en mi actividad docente me ha permitido alcanzar los resultados esperados en mis planificaciones educativas.</p>
<p>B12) Conozco el modelo educativo que regula la incorporación de las TIC en la universidad</p> <p>Notas:</p> <p><i>S4: Establece una relación con la práctica docente que sea más identificable como el plan de clase.</i></p> <p><i>S5: ¿Para qué le sirve al docente conocerlo? Mejor pregunta si lo usa de forma pragmática, como para elaborar sus planeaciones</i></p>	<p>B12) Al hacer mi plan de trabajo docente considero los lineamientos del modelo educativo que regula la utilización de las TIC en la universidad</p>
<p>B21) Comparto la cultura de incorporación de las TIC en la universidad</p> <p><i>S3: ¿Cómo? da un ejemplo</i></p> <p><i>S1: Igual, "colabora", no "comparto".</i></p>	<p>B21) Colaboro con la cultura de utilización de las TIC de la universidad compartiendo los resultados de su aplicación en reuniones de academia.</p>
<p>B22) La incorporación de herramientas tecnológicas en mi práctica docente corresponde con las lineamientos del modelo educativo de la universidad</p> <p><i>S4: Se puede unificar con la B12.</i></p>	<p>Se elimina este ítem al ya estar considerado en el ítem B12</p>
<p>C11) Uso de forma cotidiana dispositivos digitales como laptops, o tablets en la actividad docente</p> <p><i>S1: Agregar dispositivos como computadora, proyector, pizarrón electrónico, o puedes pasar este ítem como parte de las preguntas iniciales del cuestionario ¿cuáles de los siguientes dispositivos utiliza de manera cotidiana en su clase? Puede seleccionar más de una opción</i></p> <p><i>S2: Puede usar otros, ¿conviene que aquí los enlistes?</i></p> <p><i>S6: También se pueden usar Smartphone</i></p> <p><i>S3: No son los únicos que pueden usar los docentes</i></p>	<p>Este ítem se elimina de la sección de ítems en escala Likert y se incluye en la sección de datos generales del cuestionario.</p>
<p>C13) Identifico las características comunes de herramientas tecnológicas como menús de navegación opciones y configuraciones</p> <p><i>S1: Nada que ver con la dimensión si la dejas tan abierta, identifico (...) ¿para qué?</i></p> <p><i>S2: Esta pregunta es tan básica que no tiene una relación tan directa con la variable ¿de qué me sirve identificar estas características?</i></p>	<p>C13) Al utilizar las TIC en mi actividad docente puedo identificar sus componentes, como menús de navegación y configuraciones</p>
<p>C21) Uso correo electrónico para comunicarme con mis compañeros y alumnos</p> <p><i>S3: Se pueden mencionar otras formas como redes sociales, Whats app, etc. en la entrevista puedes profundizar en esto</i></p> <p><i>S4: Usar el correo, verificar el tiempo en los verbos que no sugieran una respuesta de si o no, que midan una actitud.</i></p> <p><i>S1: Lo mismo que la C11, te recomiendo que la pases al inicio del cuestionario</i></p>	<p>C21) En mi actividad docente utilizo herramientas comunicativas asíncronas como el correo electrónico para comunicarme con mis compañeros y alumnos</p>

ÍTEM ORIGINAL	ÍTEM MODIFICADO
<p>C22) Uso Skype, Hangouts u otra herramienta de comunicación inmediata para estar en contacto con mis compañeros y alumnos</p> <p><i>S1: ¿Otras? ¿Cómo cuales?, viendo la pregunta anterior te sugiero que las clasifiques como herramientas de comunicación síncronas y asíncronas, quita Skype y Hangouts, déjalo abierto.</i></p>	<p>C22) En mi actividad docente utilizo herramientas comunicativas síncronas como las conferencias en audio o video para comunicarme con mis compañeros y alumnos</p>
<p>C31) Participo con mis alumnos en redes sociales virtuales como Edmodo, Educanetwork, LinkedIn</p> <p><i>S1: Lo mismo que la A32, Mide participación, ¿pero cómo se relaciona con la práctica docente?, quita participo, a ver, ¿si uso Edmodo qué? ¿Qué tipo de redes son estas? Sin son educativas quita LinkedIn</i></p>	<p>C31) Para desarrollar mi actividad docente utilizo redes sociales educativas como Edmodo, Educanetwork como medio para poner a su alcance los contenidos de la clase.</p>
<p>C32) Participo en redes profesionales virtuales como REDOLAC, Eduteka, Red Docente OEI</p> <p><i>S1 “Participo” No, ya lo había mencionado</i></p>	<p>C32) Colaboro en redes profesionales virtuales como REDOLAC, Eduteka, Red Docente OEI para compartir mis conocimientos o conocer los de otros docentes</p>
<p>C42) Diseño escenarios de comunicación virtuales como grupos en Gdrive, Google Classroom o portales educativos</p> <p><i>S1: No queda claro si es para el trabajo colaborativo que realiza el profesor con otros profesores, o para propiciar el trabajo colaborativo entre los alumnos.</i></p> <p><i>El formato de la pregunta puede cambiar, puede ser abierta (retómala en la entrevista) ¿cuáles herramientas tecnológicas utiliza para propiciar el trabajo colaborativo entre los alumnos?, un ejemplo es Gdrive. ... algo así, si das opciones quizás solo las seleccionen por compromiso en cambio si les piden que ellos te lo digan eso da mayores posibilidades de que su respuesta sea genuina.</i></p> <p><i>S6: Se pueden utilizar otras herramientas, ¿por qué limitarlos a Google? déjalo de forma general, foros, blogs, grupos, etc. luego en la entrevista les preguntas como lo hacen</i></p>	<p>C42) Para desarrollar mi actividad docente diseño escenarios de comunicación virtuales como blogs, foros o wikis, para contribuir con el aprendizaje de mis alumnos</p>
<p>C51) Utilizo bases de datos y fuentes de información para realizar actividades de investigación en mi práctica docente</p> <p><i>S1: Verificar esta pregunta, la redacción no es clara, ¿Qué son las bases de datos, Redalyc?, ¿Fuente de información de donde, de internet? ¿Para qué?</i></p> <p><i>S2: Sugiero promover en lugar de realizar</i></p> <p><i>S6: Yo añadiría “confiables”</i></p>	<p>C51) Utilizo bases de datos en línea y/o fuentes de información confiables para promover mis actividades de investigación docentes</p>
<p>C52) Busco y filtro información de internet usando las opciones de navegadores de internet como Chrome o Mozilla</p> <p><i>S1: Puedes juntarla con la anterior, a mi parecer miden lo mismo</i></p> <p><i>S2: Uso de diversos navegadores de internet (ej. Chrome, Mozilla) para buscar información</i></p> <p><i>S6: ¿Por qué se les limita a esto?, dependiendo el tipo de búsqueda es también el navegador</i></p>	<p>C52) Utilizo diversos tipos de navegadores de internet (Chrome, Mozilla, Firefox, Opera, para hacer búsquedas de información especializadas de acuerdo a mi especialidad docente</p>
<p>C61) Divulgo los resultados de mis actividades de investigación en escenarios como congresos virtuales o publicaciones en internet</p>	<p>C61) Divulgo los resultados de mis actividades de investigación en escenarios como congresos</p>

ÍTEM ORIGINAL	ÍTEM MODIFICADO
<p><i>S1: La docencia no necesariamente implica la actividad de investigación en el estricto sentido, ¿las docentes a quien diriges el cuestionario son profesores investigadores? ¿Qué tipo de publicaciones?</i></p> <p>C62) Diseño escenarios de aprendizaje virtuales como blogs y páginas web</p> <p><i>S6: Es ambiguo hablar de comunicación virtual, ¿a qué te refieres con esto?</i> <i>S1: Creo que es parecido al C42</i></p>	<p>virtuales o revistas científicas digitales en internet</p> <p>Se elimina este ítem ya que es parecido al C42</p>
<p>C71) Participo en comunidades virtuales de formación como MOOCS, cursos en línea, seminarios virtuales</p> <p><i>S1: "Participo" Si/No ¿Para qué?</i> <i>S2: Un MOOC no necesariamente es una comunidad virtual, ten cuidado con estos términos. Además, yo agregaría el objetivo de esto si es para mejorar la docencia, etc.</i></p>	<p>C71) Para contribuir con la mejora de mi formación docente participo en curso en línea abiertos como MOOC (MiradaX, Académica, México, SINADep, etc.) o seminarios virtuales en universidades.</p>
<p>C81) Diseño materiales digitales con herramientas tecnológicas interactivas como Ardora, Educaplay, eXeLearning, Prezi</p>	<p>C81) Diseño materiales digitales con herramientas tecnológicas interactivas como Ardora, Educaplay, eXeLearning, Prezi para utilizarlos como contenidos didácticos en mi actividad docente</p>
<p>C82) Combino herramientas tecnológicas en la planificación de la actividad docente tales como presentaciones con organizadores gráficos, videos con documentos, herramientas colaborativas con gestores bibliográficos, etc.</p> <p><i>S1: La pregunta se relaciona con la G43 pueden estar más cerca, preguntar primero si incorpora tic y luego qué tipo de actividades...</i> <i>S2: Esta pregunta debe ser más específica. ¿Qué objetivo tiene esta? Yo te sugiero que la valores, quizás está implícita en otras.</i></p>	<p>C82) Al desarrollar mi actividad docente articulo la utilización de diversas herramientas tales como presentaciones con organizadores gráficos o videos con documentos digitales para enriquecer los aprendizajes de los alumnos.</p>
<p>D11) Mis alumnos me demandan que incorpore las de las TIC en las actividades de aprendizaje</p>	<p>D11) Mis alumnos generalmente me demandan que utilice las TIC en las actividades de aprendizaje</p>
<p>D12) Atiendo las demandas de que mis alumnos hacer respecto a utilizar las TIC</p> <p><i>S2: ¿Cómo?</i></p>	<p>D12) Atiendo las demandas de que mis alumnos hacen respecto a utilizar las TIC modificando mi plan de clase para incorporar algunas de estas herramientas.</p>
<p>D13) Oriento a mis alumnos para incorporar las TIC en sus actividades de aprendizaje</p>	<p>D13) Puedo orientar a mis alumnos para utilizar las TIC en sus actividades de aprendizaje</p>
<p>D21) Comparto con otros docentes estrategias para incorporan las TIC en el aula</p> <p><i>S5: Esta repetida con A31, miden lo mismo ¿Además que estrategias, detallar un ejemplo.</i> <i>S6: la pregunta ya está repetida</i></p>	<p>Se elimina este ítem debido a que ya se mencionó en el A31</p>
<p>D22) Participo en proyectos de enseñanza relacionados con la utilización de las TIC</p> <p><i>S1: "Participo" Si/No</i> <i>S2: Se repetiría con la C71 si es que la modificas</i></p>	<p>Se elimina este ítem debido a que ya se mencionó en el C71</p>

ÍTEM ORIGINAL	ÍTEM MODIFICADO
<p>E11) La universidad cuenta con la infraestructura necesaria para que incorpore las TIC en mi práctica docente</p> <p><i>S1: Sugiero que la pregunta sea más situacional, mas contextual, considero que es muy general</i>  <i>S6: ¿De qué infraestructura hablas?</i></p>	<p>E11) La universidad tiene la infraestructura que necesito para utilizar las TIC en mi práctica docente</p> <p>Se incluye también en la sección de datos generales del cuestionario.</p>
<p>E12) Cuento de forma personal con dispositivos y conectividad para incorporar las TIC en mi práctica docente.</p> <p><i>S1: Igual que la anterior, sabes que, te sugiero que estas dos preguntas vayan en al principio del cuestionario para que conozcas el entorno digital en el que se desenvuelve el docente... así le quitas la ambigüedad al ítem.</i></p>	<p>E11) Yo tengo la infraestructura que necesito para utilizar las TIC en mi práctica docente</p> <p>Se incluye también en la sección de datos generales del cuestionario.</p>
<p>E21) Uso de forma cotidiana dispositivos digitales como laptops, o tablets en la actividad docente</p> <p><i>S2: Muy general, dónde está la práctica docente</i>  <i>S6: Creo que esta pregunta sobra pues es prácticamente imposible no usar la tecnología de algún modo en el mundo actual: celular, maps, computadoras, etc.</i>  <i>S5: La pregunta ya está repetida, es la C11, unificarlas</i></p>	<p>Se elimina este ítem</p>
<p>E31) Me percibo como un docente que transforma su práctica al utilizar las TIC en la enseñanza</p>	<p>E31) Me identifico como un docente que desarrolla su actividad al utilizar las TIC como parte de sus estrategias de enseñanza</p>
<p>F11) La incorporación de las TIC en la educación está relacionada con la mejora en de la práctica docente</p>	<p>F11) La utilización de las TIC en la actividad docente está relacionada con la el desarrollo de mejora de las prácticas pedagógicas</p>
<p>F12) La incorporación de las TIC en la educación pueden mejorar las estrategias de enseñanza</p> <p><i>S1: ¿A qué te refieres con mejorar?, me parece que en el apartado teórico usas la palabra “desarrollar”, sugiero que la cambies</i></p>	<p>F12) La utilización de las TIC en la actividad docente pueden desarrollar de mejor formas las estrategias de enseñanza</p>
<p>F21) Propongo a mis alumnos proyectos de aprendizaje en donde deben incorporan las herramientas tecnológicas</p>	<p>F21) En mi planeación académica incluyo propuestas de elaboración de proyectos de aprendizaje en donde los alumnos deben utilizar las TIC</p>
<p>F22) En el aula, me siento cómodo con el uso de las tecnologías en general</p>	<p>F22) En el aula, me resulta difícil utilizar las TIC para realizar una actividad docente no planificada con anterioridad</p>
<p>F31) De acuerdo con mi experiencia, incorporar en la práctica educativa el uso de TIC podría mejorar los logros académicos</p> <p><i>S1: Pregunta de manera indirecta, no uses “logro”</i>  <i>S2: Yo hablaría de buenas prácticas y experiencias, más que logro y oportunidades</i></p>	<p>F31) En relación con mi experiencia, utilizar las TIC en la actividad docente puede ayudar a desarrollar buenas prácticas pedagógicas</p>
<p>F32) Identifico en mi práctica docente problemáticas educativas que los alumnos pueden resolver usando las herramientas tecnológicas</p> <p><i>S2: La encuentro muy general</i></p>	<p>F32) Puedo identificar en mi actividad docente problemáticas educativas que los alumnos pueden resolver usando las TIC</p>

ÍTEM ORIGINAL	ÍTEM MODIFICADO
F41) Aplico métodos de diseño instruccional apoyados en herramientas tecnológicas para planear mi práctica docente <i>S6: Los profesores no son expertos en este tema y no tienen que serlo. Esto le corresponde al colegio.</i> <i>S3: El DI se realiza previo a la práctica docente, sugiero eliminar este ítem</i> <i>S1: Considero que los ítems anteriores son parte de este, además como docente no es tan pertinente el DI como concepto</i>	Se elimina este ítem
F51) La incorporación de las TIC en la práctica docente obedece a intereses políticos	F51) La utilización de las TIC en la actividad docente obedece más a intereses políticos que académicos.
F52) La incorporación de las TIC en la práctica docente obedece a intereses académicos S3: Puedes juntar estas dos en un solo ítem,	Se elimina este ítem al unificarse con el anterior
G11) La generación de conocimientos se enriquece al incorporar las TIC en la educación	G11) En la actividad docente se pueden generar conocimientos más amplios utilizando las TIC
G21) Considero que mis alumnos podrían tener un buen nivel de satisfacción durante el proceso de enseñanza-aprendizaje si está mediado por tecnologías <i>S3: Enfocarse en el resultado, la satisfacción es subjetiva</i>	G21) Considero que mis alumnos pueden alcanzar los resultados propuestos en los objetivos de aprendizaje si mi actividad docentes esta mediada por las TIC
G31) De acuerdo con la información y conocimiento con el que cuento, la práctica docente podría verse favorecida con el uso de TIC <i>S5: Información y conocimiento son cosas distintas, dejar solo un concepto y delimita el concepto</i>	G31) Los conocimientos de los alumnos que se generan por la actividad docente pueden enriquecerse si se utilizan las TIC
G32) Me satisface incorporar estrategias educativas que incluyan aplicaciones tecnológicas <i>S1: "Satisfacción" No, no buscas medir eso</i>	G32) Las estrategias educativas en donde he utilizado las TIC me han ayudado a alcanzar mejores resultados que aquellas en donde no las utilizo.
G43) Incorporo herramientas tecnológicas en mi plan de clase de acuerdo a los objetivos de aprendizaje que requieren el uso de las TIC	G43) Identifico en mi plan de clase cuales son los objetivos de aprendizaje que requieren utilizar las TIC

Fuente: Elaboración propia.

Como se puede observar, la valoración de los expertos permitió rediseñar un cuestionario para hacer medibles las variables de estudio 1) práctica docente y 2) incorporación de las TIC en la educación, por lo que la aplicación sistemática del método Delphi permitió determinar que ítems son los idóneos, asimismo, se demostró que es eficaz para identificar áreas de oportunidad en un cuestionario al construir una lista final de ítems que recoge las diferentes percepciones de expertos en métodos de investigación y TIC.

### 5.3. Redacción final y disposición de los ítems en el cuestionario

A partir de la validación por expertos se realizó la redacción definitiva de la siguiente forma: 1) aviso de confidencialidad y privacidad, 2) objetivo del cuestionario e instrucciones, 3) datos generales del docente, 4) ítems para medir la incorporación de las TIC a partir de las dimensiones de la práctica docente y 5) agradecimiento por la participación en el cuestionario. Una copia del cuestionario terminado puede revisarse en el siguiente enlace: <https://goo.gl/forms/SzQnDbmUAhWEHyA53>.

## 6. Conclusiones

Tras analizar empíricamente la propuesta de aplicación del Método Delphi se concluye que si bien en este método se pueden utilizar distintos criterios para recuperar, sintetizar y utilizar la opinión de los expertos para validar un cuestionario, siempre es indispensable marcar adecuadamente el camino de la validación para poder obtener resultados que incidan de forma favorable en la recopilación de la información.

En este sentido, se pueden distinguir dos formas de aplicación de acuerdo a las posibilidades del investigador para llevar a cabo esta tarea, la primera a través del método que llamaremos extendido (Aponte, Cardoso y Melo, 2010; Landeta 1999; Linstone y Turoff, 1975), en donde es necesario distribuir el cuestionario, obtener la respuesta de los expertos, analizar estadísticamente las respuestas y reportar el resultado del análisis n veces hasta que la experiencia o las necesidades del investigador indiquen una suficiencia en el consenso de los participantes respecto a la calidad de los ítems que forman la encuesta.

Si bien esta aplicación garantiza la calidad extrema del cuestionario, podrá requerir de la inversión de cantidades considerables de tiempo y se puede correr el riesgo de que en el camino alguno o algunos de los expertos aborten la tarea de continuar con las revisiones, por otra parte, la segunda forma de aplicación consiste en la adecuación del método utilizado por Cabero (2014), en el cual se identifican dos fases, la primera relacionada con un análisis cuantitativo para determinar desde la percepción de los expertos que ítems deben ser incluidos y excluidos del cuestionario y la segunda en donde se les proporciona el nuevo instrumento y se les solicita hacer una valoración cualitativa mediante la anotación de argumentaciones relacionadas con la calidad de la redacción de los ítems.

Lo anterior permite obtener una calidad satisfactoria del cuestionario al ser evaluado mediante procesos confirmatorios de validación en donde el experto conoce anticipadamente en cuantas fases de evaluación participará y cuáles son las tareas que deberá hacer, lo que permite la elaboración adecuada de un cuestionario basado en las ideas de Likert (Casas, 1999) y validado desde la percepción de la aplicación del Método Delphi de Cabero (2014).

Por todo ello, se puede considerar que este artículo aporta no solo evidencia empírica del uso de un método de validación, sino que ha generado conocimiento para poder aplicar de forma sintética la validación por expertos, lo que permite obtener resultados significativos con una menor inversión de tiempo, sin embargo se debe considerar la pertinencia de profundizar en el aspecto metodológico para poder incluir estrategias que permitan afinar este tipo de validaciones.

## Referencias

- Abad, F. J., Olea, J., Ponsoda, V. y García García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Aparici, R. (2011). Principios pedagógicos y comunicacionales de la educación 2.0. *La Educación, Revista Digital*, 145, 1-14.

- Aponte, G., Cardozo, M. Á. y Melo Rosas, M. (2012). Método Delphi: Aplicaciones y posibilidades en la gestión prospectiva de la investigación y desarrollo. *Revista Venezolana de Análisis de Coyuntura*, 18(1), 41-52.
- Barroso Osuna, J. y Cabero Almenara, J. (2011). *La investigación educativa en TIC*. Madrid: Síntesis.
- Blanco, N. y Alvarado, M. E. (2005). Escala de actitud hacia el proceso de investigación científico social. *Revista de Ciencias Sociales*, 11(3), 537-544.
- Cabero Almenara, J. (2014). Formación del profesorado universitario en TIC. Aplicación del método Delphi para la selección de los contenidos formativos. *Educación XXI*, 17(1), 111-131.
- Cabero Almenara, J. e Infante Moro, A. (2014). Empleo del método Delphi y su empleo en la investigación en comunicación y educación. *EDUTECH, Revista Electrónica de Tecnología Educativa*, 48.
- Cabero Almenara, J. y Barroso, J. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón*, 65(2), 25-38.
- Cabero Almenara, J., Morales Lozano, J. A., Barroso, J., Fernández Batanero, J. M., Romero Tena, R., Román, P., Llorente, M. C. y Ballesteros, C. (2008). Creación de una guía de evaluación/autoevaluación de centros de recursos universitarios de producción de tics en la enseñanza. Pixel-Bit. *Revista de Medios y Educación*, 32, 35-55.
- Casas Castañé, M. (1999). Cambio de Actitudes en Contextos Interculturales. Modificación de Prejuicios. *Revista Electrónica de Geografía y Ciencias Sociales*, 44.
- Celina, H. y Campo, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 24(4), 572-580.
- Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, 18, 205-215.
- Christie, C. A. y Barela, E. (2005). The Delphi Technique as a Method for Increasing Inclusion in the Evaluation Process. *Canadian Journal of Program Evaluation*, 20(1), 105-122.
- D'Ancona Cea, M. A. (1999). *Metodología cuantitativa. Estrategias y técnicas de investigación social*. Madrid: Síntesis.
- Domínguez, L. y Alexis, S. (2013). ¿Ítems Politémicos o Dicotómicos? Un estudio empírico con una escala unidimensional. *Revista Argentina de Ciencias del Comportamiento*, 6(3), 30-37.
- Domínguez, S., Villegas, G., Yauri, C., Mattos, E. y Ramírez, F. (2013). Propiedades psicométricas de una escala de autoeficacia para situaciones académicas en estudiantes universitarios peruanos. *Revista de Psicología*, 2(1), 27-39.
- Fierro, C., Fortoul, B. y Rosas, L. (2000). *Transformando la práctica docente, una propuesta basada en la investigación acción*. Maestros y Enseñanza. Ciudad de México: Paidós.
- García, L. y Fernández, S. J. (2008). Procedimiento de aplicación del trabajo creativo en grupo de expertos. *Energética*, 29(2), 46-50.
- George, D. y Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*. Boston, MA: Allyn & Bacon.
- Gil Gómez de Liaño, B. y Pascual Ezama, D. (2012). La metodología Delphi como técnica de estudio de la validez de contenido. *Anales de Psicología*, 28(3), 1011-1020.
- Havelock E. (1963). *Preface to Platon*. Nueva York, NY: Harvard University Press

- Hernández Méndez, G. (2011). *Práctica docente más allá de cuatro paredes, pizarrón y mesabancos*. Veracruz: Arana Editores.
- Jay, R. y Swerdlik, M. E. (2006). *Pruebas y evaluación psicológicas*. Ciudad de México: McGraw-Hill Interamericana.
- Kerlinger, F. N. y Howard, L. B. (2002). *Investigación del comportamiento, Métodos de investigación en Ciencias Sociales*. Ciudad de México: Mc Graw Hill
- Landeta, J. (1999). *El método Delphi*. Barcelona: Ariel.
- Lee, J. y Paek, I. (2014). In Search of the Optimal Number of Response Categories in a Rating Scale. *Journal of Psychoeducational Assessment*, 32(7), 663-673.
- Linstone, H. y Turoff, M. (1975). *The Delphi Method: Techniques and applications*. Reading, MA: Addison-Wesley.
- López González, R. y Ramírez Martinell, A. (2016). Políticas de equipamiento tecnológico en Educación. *Debate Universitario*, 5(9). Recuperado de <http://portalrevisciencia.uai.edu.ar/ojs/index.php/debate-universitario/article/view/v5n9a04>
- Lozano, L. García-Cueto, E. y Muñoz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *European Journal of Research Methods for the Behavioral and Social Sciences*, 4, 73-79.
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura UNESCO. (2011). *UNESCO ICT Competency framework for teachers*. Recuperado de <http://unesdoc.unesco.org/images/0021/002134/213475E.pdf>
- Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2006). *are students ready for a technology-rich world? What PIS tells us*. París: OCDE.
- Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2016). *Education at a Glance 2016: OECD Indicators*. París: OCDE.
- Preston C. C. y Colman, A. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1-15.
- Romero, R., Cabero, J., Llorente, M. C y Vázquez Martínez, A. (2012). El método Delphi y la formación del profesorado en TIC. *Global*, 9(44), 81-93.
- Rosas Huerta, A., Sánchez Robles, J. y Chávez Cortes, M. M. (2012). La técnica Delphi y el análisis de la capacidad institucional de gobiernos locales que atienden el cambio climático. *Política y Cultura*, 38, 165-194.
- Silva, S. A., Galindo, Y. y Mendoza, L. E. (2002). Nivel de Satisfacción del Usuario de las Carreteras: Método Delphi. *Acta Universitaria*, 12(3), 41-55.
- Sireci G., S. (1998). The construct of content validity. *Social Indicators Research*, 45, 83-117.
- Trujillo, L. (2014). *El Hombre y sus medios: una aproximación al impacto mediático de la persona*. Ciudad de México: Porrúa.
- Valdés, A. A., Arreola, C. G., Angulo, J., Carlos, E. A. y García López, R. I. (2011). Actitudes de docentes de educación básica hacia las TIC. *Magis. Revista Internacional de Investigación en Educación*, 3(6), 379-392.

Valencia, T., Serna, A., Ochoa, S., Caicedo, A., Montes, J. y Chávez, J. (2016). *Competencias y estándares TIC desde la dimensión pedagógica: Una perspectiva desde los niveles de apropiación de las TIC en la práctica educativa docente*. Cali: Pontificia Universidad Javeriana.

## Breve CV de los autores

### **Carlos Enrique George Reyes**

Candidato a Doctor en Ciencias de la Educación por la Universidad Autónoma del Estado de Hidalgo con la tesis Categorización de la incorporación de las TIC en la práctica docente: Estudio comparado en universidades de Cuba y México. Maestro en Educación por la Universidad Interamericana para el Desarrollo. Consorcio Educativo Anáhuac con la tesis Diseño curricular en escenarios virtuales de aprendizaje: Caso de estudio en una universidad privada. Licenciado en Computación por la Universidad Autónoma del Estado de Hidalgo. Es profesor de posgrado en la Universidad Autónoma del Estado de Hidalgo y diversad universidades particulares, así como consultor en Educación a Distancia. Sus líneas de investigación incluyen temas relativos a la política educativa, educación comparada, uso de las TIC en la educación y competencias investigativas con TIC. ORCID ID: 0000-0002-2529-9155. Email: cgeorgemx@gmail.com

### **Laura Trujillo Liñan**

Doctora en Historia del Pensamiento por la Universidad Panamericana con la tesis Introducción al pensamiento de Marshall McLuhan desde un enfoque aristotélico. Maestra en Filosofía por la Universidad Nacional Autónoma de México (UNAM) con la tesis Primer acercamiento lógico al tema de la sustancia. Licenciada en Filosofía por la Universidad Panamericana con la tesis El acto y la potencia y su relación con la generación según el libro IX de la Metafísica. Es profesora/investigadora y Coordinadora General de la Universidad Abierta de la Universidad Panamericana. Sus líneas de investigación incluyen temas relativos a la filosofía de la tecnología, ecología de los medios, metafísica aristotélica y comunicación. Miembro del Sistema Nacional de Investigadores como Candidato, perteneciente al Consejo Nacional de Ciencia y Tecnología (CONACyT) de México, a partir de enero del 2016. ORCID ID: 0000-0003-2624-4516. Email: ltrujill@up.edu.mx

## La Experiencia de la Evaluación Docente en México: Análisis Crítico de la Imposición del Servicio Profesional Docente

### The Experience of Teacher Evaluation in México: Critical Analysis of the Imposition of the Professional Teaching Service

Francisco Guzmán Marín \*

Universidad Pedagógica Nacional (UPN) - México

La Evaluación Educativa representa una cultura, práctica y política institucional bastante tardía en el avance y fortalecimiento de la educación formal en México, pues, antes de la década de los 80's era casi inexistente en nuestro país; mientras que, por su parte, la Evaluación del Desempeño Docente en la educación obligatoria, constituye una de las áreas con menor grado de desarrollo en el Sistema Educativo Nacional, con respecto a la amplia experiencia generada en el contexto internacional. En las últimas tres décadas, el gobierno mexicano se ha empeñado en imponer la evaluación docente como una estrategia política que comprende tres principales objetivos, a saber: recuperar la rectoría del Estado en materia educativa, elevar la calidad de los servicios públicos de educación y establecer dispositivos administrativo-laborales de control sobre el ingreso, permanencia, promoción y reconocimiento en el Servicio Profesional Docente. El presente artículo expone un primer avance de investigación histórico-educativa, desde la perspectiva del enfoque genealógico nietzscheano-foucaultiano, sobre la experiencia de la evaluación docente construida en México, cuyos primeros resultados muestran ya que las políticas, instancias e instrumentos de evaluación del magisterio, impuestas por las autoridades educativas de las últimas tres décadas, no sólo resultan excesivamente onerosas sino que, además, son inconsistentes frente a los modelos y estrategias de evaluación implementadas por los sistemas educativos con mayor rendimiento académico e impertinentes al propósito manifiesto de elevar la calidad de los servicios de educación en el país.

**Palabras Clave:** Educación; Evaluación docente; Profesores; Reforma educativa; Calidad educativa.

The Educational Evaluation represents a rather late culture, practice and institutional policy in the advancement and strengthening of formal education in Mexico, since, before the Decade of the 80's, it was almost non-existent in our country; while, furthermore, the Evaluation of Teaching Performance in compulsory education, is one of the 135istó with lower degree of development in the national education system, with respect to the experience generated in the international context. In the last three decades, the Mexican Government has endeavored to impose the teaching evaluation as a political strategy that includes three main objectives, namely: recover the rectory of the State in education, raise the quality of services public education and to establish administrative control devices on admission, permanence, promotion and recognition devices in the Professional Educational Service. This article exposes a first advance of 135istórico-educativa research, from the perspective of the genealogical approach Nietzschean-Foucauldian, on the experience of teaching evaluation built in Mexico, whose first results show that policies, instances and instruments of teacher evaluation, imposed by the educational authorities of the past three decades, they are not only excessively onerous, but that they are also inconsistent with models and strategies of evaluation implemented by educational systems with higher academic performance and impertinent to the manifest purpose of raising the quality of education in the country services.

**Keywords:** Education; Teacher evaluation; Teachers; Educational reform; Educational quality.

---

\*Contacto: coraxthelastone@hotmail.com

issn: 1989-0397

www.rinace.net/riee/

https://revistas.uam.es/riee

Recibido: 31 de enero de 2018

1ª Evaluación: 23 de febrero de 2018

Aceptado: 13 de marzo de 2018

*“Es claro que la evaluación docente por sí misma no puede garantizar que un país cuente con profesores bien preparados, si antes no han existido procesos robustos de formación inicial y de formación continua.”*

(Eduardo Backhoff Escudero / Gilberto Guevara Niebla)

## 1. Introducción General

La Evaluación Educativa representa una cultura, práctica y política institucional bastante tardía en el avance y fortalecimiento de la educación formal en México, pues, antes de la década de los 80's era casi inexistente en nuestro país, como bien apunta Vidal (2009); mientras que, por su parte, la Evaluación del Desempeño Docente en la educación obligatoria, constituye una de las áreas con menor grado de desarrollo en el Sistema Educativo Nacional, a decir de Garduño (2005), con respecto a la amplia experiencia internacional, puesto que en otras naciones y regiones del mundo se viene realizando con diferentes marcos de referencia y niveles de logro, a lo largo del siglo XX –tal como sucede en los Estados Unidos de América, Cuba y Europa, por ejemplo; y aún antes del predominio del denominado neo-liberalismo en el orden global, hacia el año de 1917 en la república caribeña, de acuerdo con Valdés (2001)<sup>1</sup>–; sin embargo, ambos procesos, pero, sobre todo el último, significan factores emergentes de un vasto debate, desconfianza, escisión y resistencia empecinada, en la compleja comunidad socio-educativa mexicana, sobre todo porque los responsables directos de su diseño, arraigo y aplicación, desde la Dirección General de Evaluación (DGE), en 1984, hasta el Instituto Nacional para la Evaluación de la Educación (INEE) y la propia Secretaría de Educación Pública (SEP), en los albores del siglo XXI, han carecido de la sensibilidad y habilidad política necesaria para su instauración estandarizada y general, en la diversa composición de la estructura educativa del país.

De hecho, a tres lustros de promover la evaluación sistemática de los docentes de Educación Básica y Media Superior (EMS), y después de la experiencia construida por los recientes modelos precedentes en este campo –a saber: Carrera Magisterial y el Programa de Estímulos a la Calidad Docente–, con la imposición institucional del Servicio Profesional Docente (SPD), en la comunidad magisterial mexicana, todavía prevalecen profundas confusiones sobre los propósitos, el paradigma, los criterios, mecanismos operativos e implicaciones laborales de tal estrategia de evaluación, esto sin olvidar, el verdadero impacto que han tenido los resultados de evaluación docente en la pretendida calidad del servicio educativo, por mencionar sólo algunos de los aspectos más relevantes al respecto.

Ante la insuficiente y confusa información oficial respecto del modelo de evaluación del desempeño docente que aplica el INEE, su función institucional en el Sistema Educativo Nacional y las auténticas implicaciones laborales que comporta para el magisterio mexicano, la discusión entre el Gobierno Federal, los especialistas del ramo y los profesores agrupados en la disidencia sindical, se reduce a tres principales puntos, tales son: el presunto origen neoliberal de las políticas educativas gubernamentales, el carácter

---

<sup>1</sup> Héctor Valdés Veloz es Ex-director del Instituto Central de Ciencias Pedagógicas en Cuba.

educativo o laboral-administrativo de la reforma y la legitimidad de los procesos estandarizados de evaluación, soslayando el abordaje crítico-analítico sobre la pertinencia y oportunidad de instaurar las estrategias de evaluación docente como un factor definitorio de la reorientación histórica de las acciones de formación del profesorado, el mejoramiento de la calidad educativa y la regulación de los procedimientos de ingreso, permanencia y promoción al SPD, con el consecuente fortalecimiento de las estructuras institucionales que sustentan a la educación pública.

## 2. Principales Enfoques de Evaluación Docente

La evaluación docente, según sucede con todo tipo de evaluación, se encuentra determinada por el objeto, los propósitos, los dispositivos procedimentales, los instrumentos, los momentos y los agentes de ponderación efectiva, traducido eso en términos de interrogación, las cuestiones son: *¿qué se evalúa?, ¿para qué se evalúa?, ¿cómo se evalúa?, ¿con qué se evalúa?, ¿cuándo se evalúa? y ¿quiénes evalúan?*; de manera general, a su vez, Stake (2006) plantea como preguntas centrales: *¿cuál es la calidad del evaluando? y ¿cuáles son las principales percepciones que se tienen de su calidad?* La forma de abordar cada una de estas interpelaciones y el propio conjunto de respuestas que se proyectan al respecto, constituyen los distintos enfoques de evaluación, predominantes en los Sistemas Educativos del mundo contemporáneo. Es cierto, como bien advierte Moreno (2011), que nos encontramos en una “época marcada por constantes evaluaciones” (p. 116), se demanda que todo sea evaluado de forma permanente, ya en cuanto estrategia de fiscalización y transparencia pública, o bien en tanto instrumento de control administrativo y fuente de identificación de las áreas de mejora de cualquier fenómeno social, pero, en razón de esto mismo, es verdad, también, que el diseño, aplicación, procesamiento e interpretación de los resultados en los procesos de ponderación, pueden deformar la naturaleza y percepción del fenómeno evaluado. De hecho, Hanson (1994) considera que por su misma existencia, las pruebas modifican e incluso crean lo que se pretende evaluar. No es fortuito que en más de las ocasiones deseadas y aceptadas por los agentes evaluadores, los resultados de la evaluación tan sólo reflejen las insuficiencias de la misma, más que del objeto evaluado.

En esta perspectiva, a su vez, el modo específico en que se dispongan los dispositivos institucionales de evaluación y el uso concreto que se hagan de sus resultados, la puede convertir en un factor estratégico del desarrollo del fenómeno valorado, según parece suceder con la evaluación docente en los sistemas con mayor desempeño educativo, pero, también al contrario, es posible emplazarla en cuanto mecanismo de control y/o coerción socio-política y administrativo-laboral, como acusan los detractores del SPD, en México; cuestión ya anticipada por Beck (2007) y Landínez (2011), a los cuales se suma las reflexiones de Valdés (2001), quien sin ambages declara:

*La evaluación puede utilizarse para impulsar la realización profesional, la autonomía y la colaboración entre los docentes, o bien puede invertirse y promover recelos, miedos y rechazo expreso del profesorado debido a las desviaciones de que pueda ser objeto la evaluación y sus consecuencias para los docentes. (p. 2)*

En esta misma perspectiva, resulta pertinente señalar, siguiendo a Santos Guerra (2003a, 2003b), que la evaluación siempre comporta un factor moral, definida por los valores a que responde, los agentes beneficiados y las causas a las cuales sirve –“La evaluación, además de ser un proceso técnico es un fenómeno moral. Importa mucho saber a qué valores sirve

y a qué personas beneficia. Importa mucho evaluar bien, pero importa más saber a qué causas sirve la evaluación”, según explica el especialista español (p. 11)–. De ahí, pues, la importancia fundamental del enfoque y modelo de la evaluación docente que se pretende adoptar en un país, cuyo propósito nodal es el impulso al desarrollo y calidad de los procesos formativos. Así, en términos generales, siguiendo el análisis de Valdés (2001) e Montenegro (2007), es posible identificar cinco principales modelos de evaluación docente, los cuales son:

- *Modelo centrado en el Perfil Profesional.* Define competencias profesionales, rasgos personales y/o estándares de desempeño deseables, conforme al perfil docente de un profesor ideal, proyectado por especialistas en la materia o por principios pedagógicos de la teoría educativa, y convertidos en disposición institucional. El «perfil del docente del siglo XXI» de la UNESCO, las «diez nuevas competencias para enseñar», de Philippe Perrenoud, y las «ocho competencias docentes» de la Reforma Integral de la Educación Media Superior (RIEMS), por ejemplo.
- *Modelo centrado en los Resultados.* El énfasis se concentra en la cantidad y calidad del logro de los aprendizajes alcanzados por sus alumnos, esto es: *¿qué aprendieron?, ¿cuánto aprendieron? y ¿qué tan bien lo aprendieron?* Es un modelo de evaluación indirecta del docente. El principio fundamental de este tipo de evaluación lo constituye el reconocimiento de que la función sustantiva de cualquier profesor es la transmisión, enseñanza y/o facilitamiento de los conocimientos propuestos en los planes y programas de estudio, por ende, la ponderación se dirige hacia el grado de consecución de su aprendizaje en los discentes. Verbigracia, la ponderación del *Texas Assessment of Academic Skills* (TAAS), en Estados Unidos, la valoración de la calidad de los mentores que se deriva de los resultados de la prueba PISA y la *Evaluación de las Competencias Básicas* de los estudiantes de Colombia.
- *Modelo centrado en el Comportamiento Docente.* La evaluación se dirige hacia el conjunto de actitudes, aptitudes, disposiciones, acciones, recursos y tiempos que el profesor compromete en el desarrollo cotidiano de sus prácticas docentes. El aspecto nodal aquí es la valoración del tipo de esfuerzo y actuación que el maestro involucra en el proceso de apropiación de los aprendizajes de sus alumnos. El riesgo posible de este modelo es la sobre-valoración del activismo docente, desarticulado de los efectos reales de su labor magisterial, según advierte Montenegro (2007).
- *Modelo centrado en la Práctica Reflexiva.* La intención central de este tipo de evaluación es que el profesor reflexione, de manera metódica, sobre sus prácticas docentes, a fin de identificar con claridad sus áreas estratégicas de mejoramiento. El enfoque básico del modelo es la auto-evaluación docente y le apuesta al compromiso profesional del mentor para diseñar e instrumentar acciones concretas de transformación de su quehacer formativo, dentro de un proceso sistemático de mejora; sin embargo, al decir de Montenegro (2007), se “carece de los parámetros objetivos para determinar qué tan eficiente es su labor y qué tanto es su mejoramiento” (p. 32).

- *Modelo Integrado.* A partir del análisis de los modelos anteriores, este tipo de evaluación define cuatro campos prioritarios de la actuación docente: tales son: el primero, lo constituye el “docente mismo como persona” (Montenegro, 2007, p. 47); el segundo, es el espacio propio donde se desarrolla la relación pedagógica; el tercero, lo representa la institución educativa; y el cuarto, se conforma por el entorno del proceso educativo. En cada uno de estos campos se definen los aspectos básicos de la gestión docente para evaluar.

Por su parte, Stake (2006), más allá de otras propuestas, criterios y disposiciones, reconoce dos principales tipos de evaluación, los cuales, grosso modo, son:

- *Evaluación basada en Estándares.* Pretende evaluar la “calidad” de un proceso mediante su medición cuantitativa, con el propósito de identificar “objetivamente” los “méritos” y áreas de oportunidad del objeto evaluado, en este caso, el desempeño de los profesores.
- *Evaluación Comprensiva.* Es interpretación “personal” a partir de la recuperación de la experiencia de los agentes involucrados –profesores, alumnos, directivos, padres de familia–, las preocupaciones que median sus prácticas socio-educativas, los problemas en que se desarrollan y los significados que construyen.

Es pertinente advertir, sin embargo, que Stake (2006) considera importante la conjugación de las visiones, los procedimientos, resultados e interpretación de ambos tipos de evaluación, pues:

*En toda reflexión que hagamos contaremos con la ayuda del pensamiento basado en criterios y del interpretativo. Cada vez que evaluemos, necesitaremos tanto una evaluación basada en estándares como una evaluación comprensiva. Juntas, rara vez se suman la una a la otra, pero gracias a las visiones separadas que nos proporcionan, podemos movernos con una mayor sensibilidad, pensar con mayor profundidad, informar con mayor cautela y comprometernos más a fondo con la tarea de representar la calidad. (p. 382)*

El problema nuclear que evidencian todos estos diferentes modelos de evaluación docente es la inexplicable descontextualización de las acciones formativas realizadas por los profesores, es decir, analizan su perfil profesional sin considerar que tal es el resultado de las estrategias de formación docente, inicial y continua, instrumentadas por el mismo Sistema Educativo; los logros de aprendizaje de sus alumnos, desestimando las condiciones económico-ambientales en donde se generan; su comportamiento didáctico-pedagógico y profesional en el aula, soslayando las presiones institucionales, sindicales y sociales a las cuales se haya sometido; la auto-reflexión respecto del mejoramiento de su propio desempeño docente, pero, sin reparar en los referentes teórico-prácticos que las dinámicas del entorno institucional le posibilitan; y los distintos campos en donde se concretan las diversas disposiciones magisteriales, aunque desarticuladas de la emergencia histórico-cultural en que acontecen. Ciertamente, el Modelo Integrado plantea como campo de evaluación el entorno del proceso educativo, pero sólo en la dirección de las acciones trascendentes del marco escolar, promovidas por los mentores en cuanto “contribución al desarrollo social y cultural” del contexto, según explica Montenegro (2007, p. 54), pero, ignora las trascendentes influencias que ejerce el ambiente socio-cultural y político-económico en la resolución de las dinámicas escolares de formación.

En México, el bajo desempeño de los profesores de Educación Básica tan sólo es el producto directo de la emergente confluencia de una serie diversa de factores, entre los cuales se encuentran: la insuficiente formación profesional docente que se imparte en las Escuelas Normales, según describen explícitamente los estudios del INEE (Santos del Real y Delgado Santoveña, 2015); las acciones de resistencia político-sindicales impulsadas desde el esquema de la ideología del “maestro militante”, tanto a nivel de los estudiantes normalistas, como en la dimensión de los mentores en servicio, que se traduce en el rechazo permanente a las transformaciones curriculares, estrategias y recursos de las reformas educativas promovidas por la SEP; y los deficientes procesos de selección al ingreso y promoción en las instituciones formadoras de los profesionales de la educación, así como del reclutamiento institucional a la carrera docente, esto por mencionar algunas de los principales variables al respecto. El desempeño docente, tanto como el fenómeno educativo mismo, no se realizan en el vacío socio-histórico, en el limbo político-cultural, en la oquedad étnico-económica, por el contrario, son determinados por las circunstancias condicionantes de la época y del contexto histórico en que se desarrollan, razón por lo cual, son factores imprescindibles a considerar en sus procesos de evaluación. En sentido estricto, la evaluación docente comporta la evaluación del propio Sistema Educativo que le evalúa.

Empero, pese a la importancia del impacto de estos diversos factores en la resolución del fenómeno educativo contemporáneo y, por ende, del desempeño docente, a causa del consenso generalizado que parece persistir entre los distintos organismos, especialistas y estudios del ámbito internacional, respecto de que “son los docentes y no las escuelas quienes representan el factor institucional más importante que afecta el aprendizaje de los estudiantes” (Hill, 2016, p. s/f), y aún más, en esta misma tendencia, “un currículum es bueno en la medida que cuente con docentes buenos”, según sentencia Antonio Fuguet Smith (2006, p. 108), de acuerdo con Bruns y Luque (2014, p. 1), “los gobiernos de todo el mundo, de manera creciente, han puesto bajo la lupa su perfil profesional y estrategias de enseñanza”, de donde devienen los afanes institucionales por desarrollar modelos de evaluación docente que posibiliten, al menos, tres acciones fundamentales, estas son: la decantación sistemática del magisterio, la gestión administrativa del arraigo de las políticas que comportan las reformas educativas y, en consecuencia, la elevación de la calidad de los servicios de educación, conforme a los indicadores y estándares prevalecientes en el orden global. El peso específico que los gobiernos le otorgan a cada uno de los factores involucrados y el modo como les articulan en la definición de las políticas educativas, decide el carácter propio del modelo de evaluación docente, diseñado e implantado en cada país.

### **3. Antecedentes de la Evaluación Docente en México**

El gobierno mexicano no puede sustraerse a suscribir, producto de la consecuente presión de las diversas instancias internacionales, estas dos tendencias que definen el diseño actual de las políticas educativas, en el orden de las sociedades del siglo XXI, a saber: el recurso institucional de los dispositivos de evaluación y la sistemática depuración, formativo-laboral, del magisterio, como factores estratégicos del fortalecimiento de la calidad de la educación formalizada. La articulación emergente de ambas tendencias sustenta la determinación gubernamental de evaluar el desempeño docente, potencial en los futuros

profesores que se forman y como ejercicio profesional en los mentores ya en servicio; primero, en cuanto ámbito inherente de la expansión de los procesos de evaluación integral del Sistema Educativo –tras haber incursionado ya en la evaluación de los aprendizajes, como también en la evaluación de las escuelas– y segundo, en tanto instrumento oficial que oriente y fundamente la toma de decisiones correspondiente al ingreso, permanencia y promoción de los maestros, en el SPD de México. El desarrollo de la evaluación docente en la Educación Básica de nuestro país, se encuentra asociado íntimamente con la construcción de la experiencia en el campo de la evaluación educativa, y de manera particular, con la evaluación del aprendizaje de los alumnos.

El reconocimiento de la importancia trascendental del profesor en el centro mismo de la disfunción o del éxito de todo un sistema y/o modelo de educación, según advierte Darling-Hammond (2000), predominante en la comprensión socio-educativa contemporánea, durante las últimas tres décadas, no es un fenómeno reciente, para ser precisos, pues, sus fuentes de procedencia se remontan hasta la construcción del primer “Cuestionario de Evaluación de la Docencia por los Alumnos” (CEDA, o TESQ, por sus siglas en inglés: *Teacher Evaluation by Students Questionnaire*), de Herman Remmers, hacia 1927, para institucionalizarse con el movimiento de rendición de cuentas (*accountability*), donde se asocia la responsabilidad docente con el logro de los objetivos propuestos, en la década de los años 60’s, de acuerdo con Graciela Cordero Arroyo, et. al. (2016), y se sustenta en las diversas investigaciones desarrolladas durante los últimos seis lustros del siglo XX, en los EUA, como bien documenta Darling-Hammond (2000, p. 32), llegando a la conclusión de que “when aggregated at the state level, teacher quality variables appear to be more strongly related to student achievement than class sizes, overall spending levels, teacher salaries (at least when unadjusted for cost of living differentials), or such factors as the statewide proportion of staff who are teachers”. Actualmente, el CEDA se aplica en alrededor del 90% de las instituciones educativas de este país, al decir de García (2005).

La influencia de la experiencia norteamericana provoca que en la Educación Superior se originen los primeros estudios y propuestas de evaluación docente, en nuestro país, por lo menos dos décadas antes que se instauren los proyectos iniciales en la Educación Básica, como las investigaciones al respecto, de la Universidad Nacional Autónoma de México (UNAM), en los años 70’s, y el diseño del primer CEDA mexicano, por el Dr. Ernesto Meneses Morales, de la Universidad Iberoamericana (UIA), aproximadamente hacia el año de 1971, siguiendo la explicación de García Garduño (2005). A partir del análisis factorial de Milton Hildebrand y colaboradores (1973, p. 41), sobre “las características y capacidades de los profesores efectivos”, Meneses Morales construye su modelo de CEDA, en función del perfil del maestro universitario que se sustenta en los siguientes rasgos del “buen profesor”: a) Transmisión de Conocimientos, b) Comunicación de un Método, y c) Comunicación de su personalidad.

Por su parte, el INEE (s/fa), Banegas y Blanco (2006) parecen concordar con Martínez Rizo y Blanco (2010), respecto a que la experiencia en el diseño e impulso de las políticas de evaluación educativa es un fenómeno por demás reciente, apenas de alrededor de cuatro décadas, en nuestro país, pese a la temprana creación del Instituto Nacional de Psicopedagogía (INP), en el año de 1936, el cual tenía como propósito básico el estudio de los problemas educativos, donde el tema de la evaluación representa un factor insoslayable, sobre todo ante la confluencia emergente del uso estandarizado de las

pruebas psicométricas en la educación, en primer instancia, y la necesidad de reconocer el grado de avance en el aprendizaje escolar, a fin de legitimar socialmente la pertinencia del recién instaurado modelo de la “Escuela Socialista”, por el gobierno del General Lázaro Cárdenas del Río, en segunda instancia. Empero, los esfuerzos del INP no consiguen consolidarse, ni tampoco tienen continuidad institucional, en virtud de lo cual sus posibles aportaciones se pierden de la historia de la evaluación educativa en México, siguiendo el análisis histórico de Martínez Rizo (2001).

De ahí, pues, que la experiencia mexicana en el campo de la evaluación educativa, de acuerdo con Zorrilla y Fernández (2003), el INEE (s/fa), Banegas y Blanco (2006), Martínez Rizo y Blanco (2010), y Cordero, et. al. (2013), se desarrolle hacia el final del siglo XX y principios del siglo XXI, en tres sucesivas etapas, a saber: la primera comprende el periodo de las décadas de los años 70’s y 80’s, caracterizada por el interés manifiesto de la SEP, en la construcción de la “estadística educativa” nacional, para lo cual se mejora y sistematiza la información censal al respecto; además de que se crea el Departamento de Estudios Cualitativos de la Educación, en el año de 1972, quien realiza el primer estudio sobre la aptitud para el aprendizaje, en los púberes egresados de la Educación Primaria, del área metropolitana, de la Ciudad de México; este departamento se transforma, luego, en la Subdirección de Evaluación y Acreditación, adscrita a la Dirección General de Planeación Educativa, en 1974, instancia que organiza la primera evaluación de cobertura nacional, con el “Examen de Ingreso a la Secundaria”, cuya calificación se realiza en los propios centros escolares, sin considerar todavía la integración de algún tipo de diagnóstico nacional; después, hacia el año de 1984, la subdirección es convertida en la Dirección General de Evaluación (DGE), espacio institucional donde se diseñan y aplican las primeras “pruebas científicas” de rendimiento escolar, en muestras nacionales, según explica García Garduño (2005).<sup>2</sup>

La primera evaluación de aprendizajes, de acuerdo con Banegas y Blanco (2006), se realiza en el sexenio de José López Portillo, 1976-1982, con el proyecto *Evaluación del Rendimiento Académico de los Alumnos de 4º y 5º Grados de Educación Primaria*, mediante muestras representativas de alumnos, en el ámbito nacional. Para el siguiente sexenio, 1982-1988, el programa sectorial del gobierno de Miguel de la Madrid Hurtado atribuye a los procesos evaluativos la función de determinar el nivel de aprovechamiento escolar, en cada uno de los servicios educativos del país, para lo cual se propone el desarrollo de un “modelo integral de evaluación” que mejore significativamente sus criterios y procedimientos institucionales de operación. En este sentido, la DGE concreta su participación en la Educación Básica y Normal, a través de tres acciones sustantivas, tales son: la primera evaluación de los diez grados de la formación básica –tercero de preescolar, los seis niveles de primaria y los tres grados de secundaria–, la aplicación del “Examen de Ingreso a la Educación Normal” (EIEN) de sostenimiento federal –en 1984, como producto de la elevación formal de los estudios de educación normalista al grado de licenciatura–, y del examen de oposición a los aspirantes de plaza federal, egresados de las normales estatales.

---

<sup>2</sup> Al respecto, es pertinente apuntar que en el año de 1989, la dependencia fue renombrada como Dirección General de Evaluación, Incorporación y Revalidación (DGEIR); después, en 1994, volvió a su nombre anterior, DGE, y para el año 2005 pasa a denominarse ahora como Dirección General de Evaluación de Políticas (DGEP).

En 1989, se diseña y aplica a los alumnos de 6° grado de primaria, el *Instrumento para el Diagnóstico de Alumnos de Nuevo Ingreso a Secundaria* (IDANIS); mientras que, hacia el mismo año, el examen de ingreso a las escuelas normales se resignifica para convertirse en *Instrumento de Diagnóstico para el Ingreso a Educación Normal* (IDCIEN); en el periodo de 1992 a 1995, se delega la instrumentación de esta prueba de evaluación a las entidades federativas y a partir de 1996, es el CENEVAL quien asume la responsabilidad del proceso, a petición de los gobiernos locales.

En el periodo de 1988 a 1994, con el propósito de contrastar el desempeño académico evidenciado durante la fase estudiantil y el desempeño docente que realizan como profesores en servicio, la DGEIR desarrolla el Estudio de Evaluación y Seguimiento de Egresados de Educación Normal. Ahora bien, un fenómeno que conviene destacar, en todo este periodo, es la secrecía de los resultados de las evaluaciones aplicadas y su nulo impacto en la definición de las políticas y programas educativos nacionales, según reconoce la misma SEP (2002a) –pues, pese “al despliegue de esfuerzos realizado, los resultados no impactaron en los responsables de la toma decisiones en materia de política educativa, como tampoco en los elaboradores del currículum ni en los administradores y prestadores directos de los servicios...” y en tanto fueron connotados como “arma política”, “los resultados de evaluación se convirtieron en información confidencial” (p. 3).

La segunda etapa se encuentra enmarcada en el lapso de 1990 al año 2000. Los alcances, intervenciones e impactos de la evaluación se potencian en el Sistema Educativo Nacional, durante todo este periodo, a partir del establecimiento de diversos propósitos institucionales, que comprenden desde la acreditación de niveles educativos y la ponderación de los aprendizajes escolares, hasta la revaloración de la función docente, con el sistema de promoción horizontal, incentivos económicos y salariales que comporta la instauración del Programa Nacional de Carrera Magisterial (PNCM), en 1992. De ahí, pues, que se desarrollen diferentes procesos y dispositivos de evaluación educativa, en dos principales ámbitos, a saber: desde el interior se continua con la aplicación del IDANIS, al cual se le suma la evaluación del Factor de Aprovechamiento Escolar, dentro del marco de Carrera Magisterial, y la evaluación del Programa para Abatir el Rezago Educativo (PARE, en 1994 ambos), el Programa de Evaluación de la Educación Primaria (EVEP, 1996) y las Pruebas de Estándares Nacionales de Español y Matemáticas (EN, 1998); mientras que desde el exterior se instrumentan las pruebas internacionales del TIMSS (1995), el *Primer Estudio Comparativo y Explicativo* (PERCE) del Laboratorio Latinoamericano para la Evaluación de la Calidad de la Educación (LLECE, 1997), de la OREAL-UNESCO, y PISA (2000).

El conjunto de estas diversas acciones de ponderación del logro escolar, en la educación pública de México, fundamenta la creación del Sistema Nacional de Evaluación Educativa (SNEE), hacia la mitad de la década de los años 90's, coordinado por la DGE, cuya finalidad principal es la consolidación y articulación de los programas que en este campo desarrollan las instancias federales y estatales.

Por su parte, aunque de carácter voluntario e individual, el Programa Nacional de Carrera Magisterial representa la primera experiencia de evaluación docente, en el nivel básico de México, y deviene del Acuerdo Nacional para la Modernización de la Educación Básica (ANMEB), suscrito entre el gobierno federal, los gobiernos estatales y el SNTE, dentro del marco de los procesos de la denominada “descentralización educativa”. Y al menos en

la declaración oficial, la evaluación de este programa se concentra en el desempeño docente y el aprovechamiento escolar de sus alumnos. Carrera Magisterial inicia sus procesos en enero de 1993, pero sus efectos se hacen retroactivos al mes de septiembre de 1992. El programa considera cinco niveles de promoción horizontal y estímulo económico (A, B, C, D y E); el estímulo se determina por la relación con las percepciones de la plaza inicial, mientras que ambos se definen por la acumulación de un máximo de 100 puntos, distribuidos en seis factores de evaluación, dependiendo de la vertiente en que participan, tales son: Antigüedad, Grado Académico, Preparación Profesional, Cursos de Actualización, Desempeño Profesional, Aprovechamiento Escolar (1ª Vertiente), Desempeño Escolar (2ª Vertiente) y Apoyo Educativo (3ª Vertiente).

La evaluación docente, en la lógica institucional de Carrera Magisterial, es secuencial, seriada y completa, es decir, los diferentes niveles sólo pueden ser alcanzados en el orden establecido, tras un periodo obligatorio de permanencia en cada uno de ellos, al mismo tiempo que los seis factores deben ponderarse en cada promoción del postulante, pues, la ausencia de cualquiera de ellos, anula su participación en el proceso respectivo, de acuerdo con la normatividad establecida para tal efecto. Sin embargo, en el comienzo de la aplicación del programa, como estrategia institucional para promover la participación de los maestros y vencer las resistencias de los escépticos, o de los disidentes, a quienes deciden suscribirse al sistema horizontal de incentivos, automáticamente se les otorga un estímulo económico y dependiendo de su grado de estudios se les promueve al nivel B o C, según acusan sus detractores.

Respecto del sistema de promoción horizontal instaurado por el Programa de Carrera Magisterial, quizás conviene advertir que no sustituye, ni desplaza, al Escalafón Vertical, también conocido como “Escalafón Tradicional”, dispositivo institucional de ascenso de una categoría inferior a otra categoría superior, de los trabajadores de base de la SEP, instituido en 1930, de acuerdo con Alberto Arnaut Salgado (2013), y cuya última actualización se realiza hacia el año de 1973, con la publicación del “Reglamento de Escalafón de los Trabajadores al Servicio de la Secretaría de Educación Pública”, en el Diario Oficial de la Federación (DOF);<sup>3</sup> por el contrario, ambos sistemas escalafonarios coexisten y se aplican de manera paralela. Los trabajadores al servicio público del Sistema Educativo Nacional, se convierten en sujetos de derecho escalafonario, una vez cumplidos los seis meses de haber obtenido la asignación de una plaza inicial. El Escalafón Vertical también opera mediante la acumulación de puntos, provenientes de cuatro principales aspectos, los cuales son: Conocimiento (grado académico más superación profesional y personal), Aptitud (eficiencia e iniciativa), Antigüedad (años de servicio), Disciplina y Puntualidad.

En el ámbito de la Educación Media Superior y Superior, con el mismo propósito de promover, diagnosticar y fortalecer la calidad del desempeño y logro de los objetivos de formación de cada uno de estos niveles, mediante el recurso de los dispositivos técnico-procedimentales de la evaluación educativa, articulando las potencialidades de la

---

<sup>3</sup> La Ley de Inamovilidad de 1930 que norma la evaluación escalafonaria de los profesores de Educación Primaria, se reformula en 1933, con la Ley de Escalafón del Magisterio y posteriormente con la expedición del Reglamento de Escalafón, en 1947, siguiendo a Arnaut (2013).

evaluación interna y externa, en este mismo periodo se constituyen dos instituciones responsables de tal función, a saber: la “Comisión Nacional de Evaluación de la Educación Superior” (CONAEVA) y el “Centro Nacional de Evaluación para la Educación Superior” (CENEVAL). Las principales vertientes de evaluación que se proponen, son: la autoevaluación institucional, la evaluación externa entre pares académicos –bajo cuyo marco se constituyen, en 1991, los Comités Interinstitucionales para la evaluación de la Educación Superior (CIEES) –, y la evaluación integral del sistema de la Educación Superior. Al decir de Rodríguez, et. al. (2009), la CONAEVA, por vía de los hechos, hacia mediados de la década de los años 90’s, deja de operar “como instancia de concertación y acuerdo entre el gobierno y las universidades representadas (sic) por la ANUIES” (p. 3).

Por su parte, a recomendación de la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES), en abril de 1993, la Coordinación Nacional para la Planeación de la Educación Superior (CONPES) aprueba la constitución del Centro Nacional de Evaluación para la Educación superior (CENEVAL), en febrero de 1994, como un organismo responsable de contribuir a elevar la calidad de la educación, en el nivel del bachillerato y la Educación Superior, a través de evaluaciones externas, independientes y estandarizadas cuyo propósito nodal es la identificación del desempeño académico y el grado de dominio de los aprendizajes escolares de los jóvenes que ingresan a estos niveles educativos; al respecto es pertinente advertir que las evaluaciones del centro son complementarias, no sustitutivas, de los procesos evaluativos que desarrollan las propias instituciones, con el objeto de acreditar y/o certificar los estudios de sus alumnos.

Las líneas prioritarias de acción del CENEVAL son: por un lado, la elaboración, aplicación y sistematización de resultados de las pruebas de ingreso a la educación media superior y superior – “Examen Nacional de Ingreso a la Educación Media Superior” [EXANI-I], “Examen Nacional de Ingreso a la Educación Superior” [EXANI-II] y “Examen Nacional de Ingreso al Posgrado” [EXANI-III]–; por otro lado, el diseño, instrumentación y procesamiento de los resultados de los “Exámenes Generales para el Egreso de Licenciatura” (EGEL); y por último, la asesoría, capacitación y promoción de las experiencias de evaluación, en estos niveles de educación. En relación con la prueba EGEL es importante señalar que no evalúa conocimientos generales y/o transversales, sino saberes específicos de las diferentes opciones profesionales, de ahí que los distintos exámenes sean independientes entre sí, tanto en la formulación de las preguntas, como en los contenidos a evaluar. El examen propio de los egresados de las licenciaturas para los profesionales de la educación es el *EGEL-EDU*. En este marco, resulta conveniente señalar que en el año de 1999, la SEP emite el Acuerdo 261 mediante el cual se disponen los criterios y normas de evaluación del aprendizaje de los estudios de la licenciatura para la formación de los profesores de Educación Básica (SEP, 1999).

Un último apunte, oportuno de compartir a modo de cierre del conjunto de fenómenos definitorios de esta etapa, es el hecho de que todavía prevalece en el ánimo de las autoridades educativas mexicanas, una cierta reserva y recelo, cuando no una plena resistencia, para difundir abierta y públicamente los resultados derivados de las pruebas y exámenes de rendimiento académico, tanto nacionales como internacionales, pese al significativo impulso oficial que recibe la evaluación, en cuanto factor estratégico de los propósitos de desarrollo de la educación, en los programas sectoriales de cada gobierno del periodo y, por consecuencia, su emplazamiento en tanto recurso decisivo del esfuerzo

de promoción, fortalecimiento y mejoramiento de la calidad educativa, mediante la conformación de diversas instancias institucionales, además de la multiplicación y diversificación de los dispositivos instrumentales y/o procedimentales de evaluación.

La tercera fase del desarrollo histórico de la experiencia mexicana en el campo de la evaluación educativa, se inicia en el año 2001, de acuerdo con el INEE (s/fa), Banegas y Blanco (2006), o en el 2002, según propone Cordero, et. al. (2013). Los rasgos fundamentales de la etapa se definen por dos importantes avances institucionales, estos son: en el ámbito político, la evaluación es emplazada como un elemento imprescindible para la planeación de la educación formal y la rendición de cuentas de las autoridades educativas; mientras que en el plano estructural, se establecen los soportes normativos y organizacionales para articular, coordinar y orientar a la diversas instancias de evaluación en el país, así como de vinculación con los organismos internacionales en la materia –la OCDE, la Asociación Internacional para la Evaluación del Logro Educativo (International Association for the Evaluation of Educational Achievement, IEA) y el LLECE, por ejemplo–, siguiendo al INEE (s/fa), Banegas y Blanco (2006).

En función de estos principios rectores se crea el INEE (2002) y, por ende, se impulsa el ulterior fortalecimiento institucional de la reconstitución estratégica del SNEE, como “un conjunto orgánico y articulado de instituciones, procesos, instrumentos, acciones y demás elementos que contribuyen al cumplimiento de sus fines”, según se plantea en el Artículo 10, de la ley del instituto (DOF, 2013, p. 13), cuya integración se inicia en la década de los años 90’s. Por lo que respecta al desarrollo de la evaluación docente, en el periodo se realizan las siguientes acciones: por un lado, desde el 2004, el CENEVAL diseña e instrumenta los “Exámenes Generales e Intermedios de Conocimientos” para el cuarto y octavo semestre de la Educación Normal –los exámenes intermedios se aplican por primera vez en el 2010, de acuerdo con la SEP (s/f)–; por otro lado, el establecimiento del “Concurso Nacional para el Otorgamiento de Plazas Docentes” (2008), del “Programa de Estímulos a la Calidad Docente” (2008), del “Examen Nacional para la Actualización de los Maestros en Servicio” (ENAMS, 2009) y del Acuerdo para la “Evaluación Universal de Docentes y Directivos en Servicio de Educación Básica” (2011); y por último, la expedición de la “Ley General del Servicio Profesional Docente” (LGSPD, 2013) y la consecuente instauración del “Programa de Promoción en la Función por Incentivos en Educación Básica” (2015).

El concurso para el ingreso al servicio magisterial, los estímulos a la calidad docente, la profesionalización permanente del profesorado y la evaluación universal, derivan de los acuerdos suscritos entre el gobierno federal y el SNTE, en mayo del 2008; los cuales tienen el objeto de fomentar una profunda transformación del Sistema Educativo Nacional, que permita elevar la calidad de la educación mexicana, en el marco de la denominada Alianza por la Calidad Educativa (ACE). Los acuerdos de la ACE se sintetizan en una diversa serie de compromisos institucionales que se agrupan en cinco principales líneas de acción, a saber: primero, modernización de los centros escolares; segundo, profesionalización de los maestros y de las autoridades educativas; tercero, bienestar y desarrollo integral de los alumnos; cuarto, formación integral de los alumnos para la vida y el trabajo; y quinto, evaluar para mejorar. En el acuerdo sobre la profesionalización de los mentores se establece que el ingreso y la promoción de los docentes se realizan por vía del concurso público de oposición, además de estimular el mérito profesional de los maestros, en “función exclusiva de los resultados de logro de sus alumnos” (SNTE, 2008, p. 16);

mientras que en el compromiso relativo a la evaluación se plantea con toda claridad la articulación del Sistema Nacional de Evaluación, la “evaluación exhaustiva y periódica de todos los actores del proceso educativo” y la definición de los estándares de desempeño, en cada ámbito del fenómeno de la educación formal (SNTE, 2008, p. 24).

Así, pues, el Concurso Nacional para el Otorgamiento de Plazas Docentes se dirige a los egresados de las licenciaturas afines a la educación, que opten por el acceso a una plaza docente definitiva, adicional o, en su defecto, incremento de horas, en los niveles de la Educación Básica. En el 2008, la evaluación la realiza la DGEP, con una prueba que comprende tres áreas sustantivas: *contenidos de enseñanza, competencias didácticas y habilidades intelectuales básicas*, distribuidas en 80 reactivos. Para el 2009, conforme a los acuerdos de la ACE, se constituye el Órgano de Evaluación Independiente con Carácter Federalista (OEIF), integrado por 72 representantes de la SEP y el SNTE, que tiene como responsabilidad la dictaminación técnica y académica de los criterios de acreditación del concurso, los aspectos de las evaluaciones estatales adicionales, los reactivos de los exámenes y las guías para los sustentantes. A partir de ese año, y a solicitud de la DGEP, el CENEVAL es la instancia encargada de diseñar, coordinar y aplicar los 24 instrumentos estandarizados del “Examen Nacional de Conocimientos y Habilidades Docentes” (ENCHD).<sup>4</sup> La prueba del CENEVAL consideran cuatro tipos de dominio, tales son: *contenidos curriculares, competencias didácticas, habilidades intelectuales específicas y normatividad, gestión y ética docente*, desglosados en 80 reactivos de opción múltiple y 30 reactivos piloteados; el tiempo máximo de resolución del examen es de tres horas, todo esto siguiendo a Santiago, et. al. (2012), Nájera y Mendieta (2014). El diseño del ENCHD carece de un referente preciso sobre el perfil profesional que deben detentar quienes aspiran a incorporarse al magisterio, como bien reconocen Nájera y Mendieta (2014).

El Programa de Estímulos a la Calidad Docente, previsto en los compromisos de la ACE, por su parte, constituye un dispositivo de evaluación indirecta del desempeño magisterial, cuyo propósito básico es el reconocimiento individual y colectivo de los profesores que presentan los mejores logros de aprendizaje de sus alumnos, en las pruebas nacionales – “Evaluación Nacional de Logros Académicos en Centros Escolares” (ENLACE), verbigracia–, mediante estímulos salariales, con la intención expresa de promover las prácticas docentes que contribuyan a elevar la calidad educativa. La participación en el programa no es obligatoria y los estímulos se asignan de manera colectiva e individual.

En otro de los acuerdos de la ACE se plantea el compromiso de proporcionar una formación continua a los profesores en servicio, >>orientada hacia la calidad educativa<<, por lo que el anterior Programa Nacional para la Actualización Permanente de los Maestros de Educación Básica en Servicio (PRONAP), instaurado en 1994, se transforma en el “Sistema Nacional de Formación Continua y Superación Profesional” (SNFCSP), hacia el año 2008, el cual se acompaña de los ENAMS como un dispositivo oficial de acreditación de los dominios de aprendizaje desarrollados en los trayectos formativos de la práctica docente, diseñados por el sistema, en tres principales áreas, estas son: *el*

---

<sup>4</sup> En el año del 2009 sólo fueron 17 instrumentos diseñados, pero, para el 2010 se incrementaron seis instrumentos más y en el 2011 se diseña otro para la Educación Inicial, con lo cual se integran 24 exámenes en total.

*conocimiento de las disciplinas y de los enfoques de enseñanza, de los procesos de aprendizaje de los niños y jóvenes; el manejo de los materiales educativos; y la comprensión general de los problemas educativos y de la gestión escolar.* Una vez más el CENEVAL es la instancia responsable del diseño, validación, seguimiento de la aplicación y calificación de los exámenes. Los ENAMS no son obligatorios, son pruebas estandarizadas, están dirigidos a los profesores de Educación Básica que se desempeñen como docentes frente a grupo, en funciones directivas o como ATP, son específicos para cada curso ofertado por el sistema y tienen como consecuencia dos productos posibles, a saber: un Diploma SEP de acreditación del dominio presentado y un folleto personal de las principales áreas de mejora del sustentante, de acuerdo con SEP (2008) y Santiago y otros (2012).

En el mismo marco de la ACE, la SEP y el SNTE suscriben el Acuerdo para la Evaluación Universal de Docentes y Directivos en Servicio de Educación Básica, con un doble propósito, esto es: por un lado, construir un “diagnóstico integral” de las competencias profesionales de los docentes frente a grupo, directores y ATP, así como del logro académico de sus alumnos; y por otro lado, “focalizar los trayectos de formación continua en las áreas de oportunidad que se detecten con base en sus resultados” (SEP/SNTE, 2011, p. 2); además, se pretende que sus productos orienten la “calidad y pertinencia” de los programas de estudio de la Educación Básica y Normal, así como del diseño y aplicación de políticas educativas eficaces en la consecución de la calidad del Sistema Educativo y del aprendizaje deseable en los estudiantes.

La evaluación es obligatoria para los profesores frente grupo, en funciones directivas y de asesoría técnico-pedagógica, de las «escuelas públicas y privadas»; la periodicidad es trianual, llevándose a cabo el segundo y tercer fin de semana del mes de junio, del año correspondiente. Los resultados de la evaluación se consideran para la acreditación de los factores respectivos de Carrera Magisterial y de los Estímulos a la Calidad Docente, conforme a los lineamientos de cada programa.

La evaluación universal se acompaña de los trayectos de formación continua, principalmente en el campo de las matemáticas, el español, las ciencias, las tecnologías de la información y el inglés, que conforman el SNFCSP; sus resultados son de carácter público y en conjunto con los trayectos particulares que devienen de las áreas de oportunidad identificadas, se informan a las autoridades educativas estatales, al sindicato de los maestros, personal docente, directivo y de supervisión escolar, según explica el Acuerdo establecido entre la SEP y el SNTE. En el diseño de los instrumentos de la evaluación universal participan docentes destacados, directivos, especialistas y asesores académicos de la Educación Básica, así como el INEE, desde luego; con esto “se asegura la pertinencia de los instrumentos de evaluación y se evita el uso unilateral o tendencioso de los mismos”, de acuerdo con la SEP (2012, s/p).

Por su parte, a fin de arraigar la institucionalización de la evaluación docente en México, la Reforma Educativa, promovida por el gobierno de Enrique Peña Nieto, en la segunda década del siglo XXI, impulsa cuatro principales acciones, a saber: la consolidación de la autonomía plena del INEE como órgano responsable de coordinar la implementación de las evaluaciones nacionales e internacionales en la educación obligatoria; el fortalecimiento de la estructura organizacional del SNEE; la promulgación de la Ley General del Servicio Profesional Docente (LGSPD) que regula los procesos de ingreso, permanencia, promoción y reconocimiento de los profesores en el SPD; además de la

implementación sistemática –no exenta de críticas contingencias políticas– de un intrincado, confuso e irrelevante proceso de evaluación docente.

#### **4. Consideraciones Finales**

En principio, conviene advertir que en una acuciosa revisión de los sistemas educativos de mayor logro académico, encontramos que la evaluación docente no representa el factor nuclear del mejoramiento de la calidad educativa, como se pretende en México, sino más bien un cierto dispositivo institucional de recuperación de información estratégica para reforzar los procesos de formación y desarrollo profesional de los mentores, y por consecuencia lógica, de los servicios públicos de la educación formal. A su vez, la asociación de las acciones evaluativas con los incentivos económicos genera tres tipos de desviaciones educativas, a saber: la corrupción laboral, la simulación docente y la reducción de las prácticas formativas al desarrollo de capacidades para resolver exámenes, como bien parece advertir Aboites (2013), en el *1<sup>er</sup> Foro Regional: Análisis y Perspectivas sobre la Reforma Educativa*, realizado en la ciudad de Morelia, Michoacán.

La orientación política, la función institucional y los dispositivos procedimentales asignados por la SEP y el INEE, a la evaluación docente, más bien parecen promover el perfil profesional de un “maestro idóneo” capaz de responder a la exigencia de los trayectos evaluativos y de formar alumnos competentes para aprobar exámenes de rendimiento escolar, que a educar ciudadanos, como ya viene sucediendo en nuestro país, de acuerdo con Santiago y otros (2012): “Enseñar para la prueba se ha convertido en una práctica pedagógica generalizada en México. Los directores de escuela, maestros y alumnos consideran que practicar para las pruebas estandarizadas es la mejor estrategia para elevar el logro de los alumnos” (p. 82); incluso, ya en el marco del SPD, los mismos profesores que resultan “no idóneos”, conforme a las debilidades de desempeño en la función y/o las áreas de oportunidad en el conocimiento, evidenciadas en los exámenes de permanencia, son “preparados” para acreditar las pruebas subsecuentes, no para potenciar su desarrollo profesional.

A efectos de mejorar los procesos evaluativos de la educación, en México, se constituye un ostentoso, voluminoso y oneroso aparato rector de las estrategias institucionales de evaluación, según previene ya Carlos Ornelas (2013), apenas recién instaurada la plena autonomía del INEE –si “se cumplen esos propósitos [la estructura organizacional dispuesta en el Estatuto Orgánico propuesto por la Junta del Gobierno del instituto], en pocos años tendremos una institución gigantesca, burocrática y pesada, sin rumbo” «y sin eficacia en el desempeño de sus funciones particulares», bien es posible completar el planteamiento del reconocido investigador educativo (s/p)–, pero que, al menos en el ámbito de la evaluación docente, no ha conseguido el establecimiento de un determinado trayecto evaluativo transparente, organizado y equitativo que revalorice la labor magisterial de los profesores y corresponda realmente con las exigencias sociales de su desarrollo profesional y el consecuente mejoramiento de la calidad educativa, siguiendo las conclusiones derivadas de la consulta a los maestros evaluados en el 2015, realizada por este mismo organismo (INEE, 2016b).

El breve periodo de construcción de la experiencia mexicana en materia de evaluación educativa, en lo general, y de evaluación docente, en lo particular, se caracteriza por la

excesiva multiplicación y diversificación de las disposiciones, instancias, trayectos, dispositivos procedimentales e instrumentos de evaluación, en el Sistema Educativo Nacional, que lejos de trazar una determinada línea de perfeccionamiento progresivo de las estrategias evaluativas, más bien evidencia rupturas innecesarias, abruptas modificaciones y duplicidad de acciones, sin apenas tiempo suficiente para valorar su verdaderos alcances socio-educativos, impactos institucionales y áreas de mejoramiento técnico, a lo cual deben agregarse las propuestas internacionales de ponderación del logro educativo, como bien se puede apreciar en los cuadros 1, 2 y 3.

Cuadro 1. Dispositivos de Evaluación Educativa. 1930-1988

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
1930-1973	Escalafón Vertical	SEP-SNTE	Nacional	Promoción vertical de los profesores de Educación Básica
1971	CEDA	UIA	Universitario	Perfil de los profesores universitarios
1972	Estudio sobre la Aptitud para el Aprendizaje	Departamento de Estudios Cualitativos de la Educación	Local	Egresados de la Educación Primaria
1974	Examen de Ingreso a la Secundaria	Subdirección de Evaluación y Acreditación	Nacional	Selección de alumnos para el ingreso a la Educación Secundaria
1976	Acuerdo 3810	Consejo Nacional Técnico de la Educación (CONALTE)	Nacional	Establecimiento de los procedimientos de evaluación del aprendizaje de los alumnos de Educación Básica y Normal.
1976-1982	Evaluación del Rendimiento Académico de los Alumnos de 4° y 5° Grados de Educación Primaria	Subdirección de Evaluación y Acreditación	Nacional	Rendimiento escolar de los alumnos de 4° y 5° grados de Educación Primaria
1978	Acuerdo 17	SEP	Nacional	Establecimiento de las normas que regulan los procedimientos de evaluación del aprendizaje, en los distintos tipo y modalidades de la educación bajo el control directo de la SEP
1982-1988	Evaluación de los Diez Grados de la Educación Básica	DGE	Nacional	Rendimiento escolar de los alumnos de Educación Básica
1983	Instrumento de Diagnóstico y Clasificación para el Ingreso a Educación Normal (IDCIEN)	DGE	Nacional	Proceso de selección de aspirantes a ingresar a la Educación Normal
1984	Examen de Ingreso a la Educación	DGE	Nacional	Aspirantes a ingresar a las Escuelas Normales federales

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
1984	Normal (EIEN) Sistema Nacional de Investigadores	CONACyT	Nacional	Producción de docentes-investigadores con el propósito de otorgar apoyos económicos
1982-1988	Examen de Oposición a los Aspirantes de Plaza Federal	DGE	Nacional	Egresados de las Escuelas Normales Estatales

Fuente: Elaboración propia con base a fuentes oficiales y de investigadores educativos.

Cuadro 2. Dispositivos de Evaluación Educativa. 1989-1999

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
1989		DGE	Nacional	
1992-1995	IDCIEN	Autoridades Locales	Local	Aspirantes a ingresar a las Escuelas Normales federales
1996		CENEVAL		
1989	IDANIS	DGE	Nacional	Aspirantes de ingreso a la Educación Secundaria
1989	Comisión Nacional de Evaluación de la Educación Superior (CONAEVA)	SEP-ANUIES	Nacional	Calidad de la Educación Superior
1988-1994	Estudio de Evaluación y Seguimiento de Egresados de Educación Normal	DGE	Nacional	Contrastar el desempeño académico de la época de estudiantes normalistas con el desempeño docente frente a grupo
1992	Acuerdo 165	SEP	Nacional	Establecimiento de las normas de evaluación del aprendizaje en la Educación Primaria, Secundaria y Normal
90's	Autoevaluación Escolar	SEP-DGEP	Focal	Estrategia de mejora continua
1991	CIEES	CONPES	Nacional	Valoración inter-institucional de la calidad de los programas educativos de Educación Superior, por pares Académicos
1992	Programa Nacional de Carrera Magisterial	SEP-SNTE	Nacional	Revaloración de la función docente mediante la instauración del Sistema de Promoción Horizontal, incentivos económicos y salariales
1994	Preparación Profesional			Medición de los conocimientos y habilidades para el desarrollo de las funciones docentes.

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
	Aprovechamiento Escolar			Medición del rendimiento académico de los alumnos de 3° a 6° de Educación Primaria y de los tres grados de la Educación Secundaria, de los mentores que participan en Carrera Magisterial.
1994	Evaluación Cualitativa del PARE	Consejo Nacional de Fomento Educativo (CONAFE)	Focalizado	Estudio de unidades escolares asociadas al PARE, con el propósito de estimar los procedimientos de implantación y los cambios posibles generados por el Programa, en los Estados de Guerrero y Oaxaca
1994	Acuerdo 200	SEP	Nacional	Establecimiento de las normas de evaluación del aprendizaje en la Educación Primaria, Secundaria y Normal
1994	Exani-ii	CENEVAL	Nacional	Identificación del desempeño académico y el grado de dominio de los aprendizajes escolares de los jóvenes que ingresan a la EMS y Superior
1994	EGEL-EDU	CENEVAL	Nacional	Reconocimiento de saberes profesionales específicos de los egresados de la educación normal
1995	TIMSS	IEA	Internacional	Dominio de los conocimientos de matemáticas y ciencias de los estudiantes de cuarto y octavo grado
1996	AEE	Instancias Estatales	Local	Fortalecimiento de las tareas evaluativas en las entidades federativas
1996	Comité Interestatal Técnico de Evaluación (CTIE)	DGE	Nacional	Concertación federal y estatal de las AEE
1996	EVEP	DGE/CONAFE	Nacional	Conformación de series históricas sobre el logro educativo en la Educación Primaria y, al propio tiempo, consolidación de AEE.
1997	Evaluaciones Cualitativas	DGE	Nacional	Estudio de las dinámicas escolares de gestión educativa y práctica pedagógica, en la

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
				Educación Primaria y Secundaria, con el fin de explicar el logro académico
1997	PERCE	LLECE-OREALC/UNESCO	Internacional	Aprendizaje de los estudiantes de 3er y 4to grado de Educación Primaria, en el campo de la lectura y la matemática
2006	SERCE			Aprendizaje del lenguaje –lectura y escritura– y la matemática en
2013	TERCE			estudiantes de 3er y 6to de Educación Primaria, así como de ciencias naturales, en este último nivel
1998	Estándares Nacionales	DGE	Nacional	Medición del grado de desarrollo de las habilidades para la comprensión lectora y la resolución de problemas matemáticos
1998 2000 2003 2006 2009 2012 2015	PISA	OCDE	Internacional	Competencias básicas para la vida y el trabajo de los jóvenes de 15 años
1999	Acuerdo 261	SEP	Nacional	Criterios y normas de evaluación del aprendizaje de los estudios de licenciatura para la formación de los profesores de Educación Básica

Fuente: Elaboración propia con base a fuentes oficiales y de investigadores educativos.

Cuadro 3. Dispositivos de Evaluación Educativa. 2000-2015

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
2000	COPAES	-----	Nacional	Acreditación de órganos certificadores de programas académicos
2001	Evaluación del Programa de las Escuelas de Calidad (PEC)	DGE- Coordinación Nacional del PEC-AEE	Nacional	Valorar el impacto del PEC en los centros educativos, respecto del logro académico, las prácticas pedagógicas, la rendición de cuentas y la autoevaluación
2004	Exámenes Generales e Intermedios de Conocimientos	CENEVAL	Nacional	Diagnóstico general de la formación docente
2010				Diagnóstico intermedio de la formación docente

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
2005	Excale	INEE	Nacional	Valoración del grado en que los alumnos alcanzan los aprendizajes establecidos en los planes y programas de estudio
2006	Enlace	SEP	Nacional	Nivel de dominio en español, matemáticas y una tercera asignatura rotativa anualmente, conforme a los planes y programas oficiales de estudio, en Educación Básica; dominio de las habilidades para la vida en español, matemáticas y las competencias establecidas en la RIEMS, el bachillerato
2006	Estudios sobre las Condiciones de la Oferta Educativa (COE)	INEE	Nacional	Valoración de los recursos y procesos de las escuelas de Educación Básica
2008	Concurso Nacional para el Otorgamiento de Plazas Docentes	DGEP	Nacional	Selección de los profesionales de la educación que opten por el acceso a una plaza docente definitiva, adicional o incremento de horas en Sistema Educativo Nacional
2008	Programa de Estímulos a la Calidad Docente	SEP	Nacional	Reconocimiento de los maestros de Educación Especial y Básica cuyos alumnos presentan alto desempeño escolar
2009	Sistema Nacional de Acreditación de Centros Escolares de Educación Básica (SNACEEB)	DGEP	Nacional	Ponderación de la calidad de los ámbitos, componentes y funciones de los centros escolares de la Educación Básica
2009	ENCHD	CENEVAL	Nacional	Valoración de las habilidades docentes para la enseñanza, en la educación obligatoria
2012	ENCHCD			
2009	ENAMS	CENEVAL	Nacional	Acreditación de los dominios de aprendizaje desarrollados en los trayectos formativos de la práctica docente, diseñados por el Sistema Educativo Nacional
2011	Evaluación Universal de Docentes y Directivos en Servicio de Educación Básica	INEE	Nacional	Construcción de un diagnóstico integral de las competencias docentes frente a grupo, directivos y ATP, así como del logro académico de sus alumnos
2014	Planea	INEE	Nacional	Identificación del grado de dominio de los aprendizajes escolares, en los estudiantes de la educación obligatoria, conforme a los planes y programas de estudio
	ELSEN			Identificación del logro académico de los alumnos
	ELCE			Reconocimiento de aprendizajes clave
	EDC			Identificación de aprendizajes clave en los campos formativos del

AÑO	INSTRUMENTO O PROGRAMA	INSTITUCIÓN	ÁMBITO	OBJETO DE EVALUACIÓN
				lenguaje y comunicación, además de matemáticas
2014	Concursos de Ingreso al SPD	INEE	Nacional	Valoración de la suficiencia en el dominio de los conocimientos y capacidades para el desempeño docente, en los profesores que se incorporan al SPD
2015	Evaluación del Desempeño Docente	INEE	Nacional	Identificación de la idoneidad de los maestros en servicio
2015	Programa de Promoción en la función por Incentivos en Educación Básica	INEE	Nacional	Reconocimiento a profesores de alto desempeño y resultados sobresalientes en las evaluaciones adicionales

Fuente: Elaboración propia con base a fuentes oficiales y de investigadores educativos.

En síntesis, los diferentes dispositivos institucionales de la evaluación educativa, sobre los cuales se viene construyendo la experiencia mexicana en este campo, pueden agruparse en tres principales estrategias gubernamentales, estas son: por un lado, el establecimiento de políticas, programas y acuerdos secretariales que fundamentan, orientan y regulan los diversos procesos evaluativos, en los distintos niveles del Sistema Educativo Nacional; por otro lado, la conformación de instancias, dependencias, organismos, instituciones y sistemas de carácter federal y/o estatal, responsables del diseño, validación y aplicación de los instrumentos de evaluación, así como de la organización, procesamiento y difusión de sus resultados; y por último, la construcción de estudios, diagnósticos y recursos instrumentales de valoración del desempeño de los servicios de la educación formal, con el objeto de promover el mejoramiento de su calidad académica. Al decir de Hugo Aboites (2012), en casi 25 años se han aplicado más de 110 millones de exámenes estandarizados a niños, jóvenes, profesores y escuelas. Incluso, en esta compulsiva tendencia a evaluar todo, las propias disposiciones de evaluación también son evaluadas por diferentes agentes internacionales y nacionales, como la implementada por la *Research and Development Corporation* (RAND) al Programa de Carrera Magisterial, los análisis y valoraciones de la OREALC/UNESCO a la Evaluación del Desempeño de Docentes, Directivos y Supervisores de la Educación Básica y Media Superior, del 2015-2016, las revisiones generales de los equipos de la OCDE a los procesos evaluativos de nuestro país y el Estudio de Validación de las pruebas Enlace y Excale, realizado por el equipo que coordina Felipe Martínez Rizo (2015), por ejemplo; pero que, en su conjunto, no han logrado propiciar la elevación de la calidad de los servicios educativos en el país y mucho menos reposicionar al Sistema Educativo Nacional en el ranking de las evaluaciones internacionales.

## Referencias

- Aboites, H. (2012). La disputa por la evaluación en México: Historia y futuro. *El Cotidiano, Revista de la Realidad Mexicana Actual*, 176 (noviembre-diciembre), 1-25.
- Banegas, I. y Blanco Bosco, E. (2006) *Políticas y sistemas de evaluación educativa en México. Avances, logros y desafíos*. Ciudad de México: INEE.

- Beck, U. (2007) *Teoría de la sociedad del riesgo*. En A. Giddens et. al. (Eds.), *Las consecuencias perversas de la modernidad* (pp. 201-222). Barcelona: Anthropos.
- Bruns, B. y Luque, J. (2014). *docentes excelentes: cómo mejorar el aprendizaje en América Latina y el Caribe. Resumen*. Washington, DC: Banco Mundial.
- Cordero Arroyo, G. (2013). La evaluación docente en Educación Básica en México: panorama y agenda pendiente. *Sinéctica, Revista Electrónica de Educación*, 41.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A Review of state policy evidence. *Education Policy Analysis Archives*, 8(1),1-44.  
<https://doi.org/10.14507/epaa.v8n1.2000>
- Fuguet Smith, A. (2006) *Educación en Canadá: Protocolos y elementos de reflexión para el desarrollo curricular*. Caracas: Colecciones Estudios Canadienses.
- García Garduño, J. M. (2005). el avance de la evaluación en México y sus antecedentes. *Revista Mexicana de Investigación Educativa*, 10(27), 1275-1283.
- Hanson, F. A. (1994). *Testing testing: social consequences of the examined life*. Berkeley, CA: University of California Press.
- Hill, H. C. (2016). *Evaluación de maestros en Estados Unidos: Un caso de estudio*. Recuperado de <http://politikon.es/2016/04/19/18152/#>
- INEE. (2006). *Los cuatro primeros años. Avances y desafíos*. Ciudad de México: INEE.
- INEE. (2007). *Plan de Desarrollo del Sistema Nacional de Indicadores Educativos de México, 2007-2014*. Ciudad de México: INEE.
- INEE. (2013). *Estatuto Orgánico del Instituto Nacional para la Evaluación de la Educación*. Ciudad de México: DOF.
- INEE. (2015a). *Plan Nacional para la Evaluación de los Aprendizajes (PLANEA)*. Ciudad de México: INEE.
- INEE. (2016a). *Criterios Técnicos y de Procedimientos para el Análisis de los Instrumentos de Evaluación, el Proceso de Calificación y la Emisión de Resultados*. Ciudad de México: DOF.
- INEE. (2016b). *Evaluación del Desempeño desde la Experiencia de los Docentes. Consulta con Docentes que Participaron en la Primera Evaluación del Desempeño 2015*. Ciudad de México: INEE.
- INEE. (2017a). *La Evaluación del Desempeño, 2017*. Ciudad de México: INEE.
- INEE. (2017b). *La Educación Obligatoria en México. Informe 2017*. Ciudad de México: INEE.
- INEE. (s/fa). *Breve recorrido por la evaluación de la Educación Básica en México*. Ciudad de México: INEE.
- INEE. (s/fb). *Evaluación docente*. Ciudad de México: INEE.
- Landínez, D. (2011). Resistiendo al control. ¿Es posible una ética de la resistencia?. *Saga, Revista de Estudiantes de Filosofía*, 22, 37-46.
- Martínez Rizo, F. (2001). Evaluación educativa y pruebas estandarizadas. Elementos para Enriquecer el Debate. *Revista de la Educación Superior*, 120(30), 1-12.
- Martínez Rizo, F. (2013). El Futuro de la Evaluación Educativa. *Sinéctica, Revista Electrónica de Educación*, 40.
- Martínez Rizo, F. (2015a). *Las Pruebas EXCALE para Educación Básica. Una Evaluación para el Instituto Nacional para la Evaluación de la Educación*. Ciudad de México: INEE.

- Martínez Rizo, F. (2015b) *Las Pruebas ENLACE y EXCALE*. Ciudad de México: INEE
- Martínez Rizo, F. y Blanco, E. (2010). La Evaluación Educativa en México: Experiencias, Avances y Desafíos. En A. Arnaut y S. Giorguli (Coords.), *Los Grandes Problemas de México. VII. Educación* (pp. 89-124). Ciudad de México: El Colegio de México.
- Montenegro Aldana, I. A. (2007). *Evaluación del desempeño docente. fundamentos, modelos e instrumentos*. Bogotá: Cooperativa Editorial Magisterio.
- Moreno Olivos, T. (2011). La cultura de la evaluación y la mejora de la escuela. *Perfiles Educativos*, 23(131), 116-130.
- Santiago, P., McGregor, I., Nusche, D., Ravela, P. y Toledo, D. (2012). *Revisiones de la OCDE sobre la evaluación en educación*. Ciudad de México: OCDE.
- Santos del Real, A. y Delgado Santoveña, A. (Coords.) (2015). *Los docentes en México. Informe 2015*. Ciudad de México: INEE.
- Santos Guerra, M. A. (2002). Una flecha en la diana. la evaluación como aprendizaje. *Andalucía Educativa*, 34, 7-9.
- Santos Guerra, M. A. (2003a). Dime cómo evalúas y te diré qué tipo de profesional y de persona eres. *Revista Enfoques Educativos*, 5(1), 69-80.
- Santos Guerra, M. A. (2003b). *La evaluación como aprendizaje. Una flecha en la diana*. Madrid: Narcea.
- SEP. (1999). *Acuerdo Número 261 por el que se Establecen Criterios y Normas de Evaluación del Aprendizaje de los Estudios de Licenciatura para la Formación de Profesores de Educación Básica*. Ciudad de México: DOF.
- SEP. (2002a). *La Experiencia de la Dirección General de Evaluación en la Educación Básica y Normal. 30 Años de Medición del Logro Educativo*. Ciudad de México: SEP.
- SEP. (2002b). *Decreto por el que se Crea el Instituto Nacional para la Evaluación de la Educación*. Ciudad de México: DOF.
- SEP. (2008). *Folleto Informativo. Exámenes Nacionales para la Actualización de los Maestros en Servicio (ENAMS) La Asesoría Académica en Educación Básica*. Ciudad de México: CENEVAL.
- SEP. (2011). *OCDE Revisión de los Marcos de Valoración y de Evaluación para Mejorar los Resultados Escolares. Informe de las Prácticas de la Evaluación de la Educación Básica en México. 2010*. Ciudad de México: SEP.
- SEP. (2013a). *Lineamientos que Regulan el Programa de Estímulos a la Calidad Docente*. Ciudad de México: SEP.
- SEP. (2013b). *Ley del Instituto Nacional para la Evaluación de la Educación*. Ciudad de México: INEE.
- SEP. (2015). *Etapas, Aspectos, Métodos e Instrumentos. Proceso de Evaluación del Desempeño Docente. Educación Básica. Ciclo Escolar 2015-2016*. Ciudad de México: SEP.
- SEP. (2017). *Etapas, aspectos, métodos e instrumentos. Proceso de evaluación del desempeño docente. Educación Básica. Ciclo Escolar 2017-2018*. Ciudad de México: SEP.
- SEP/SNTE. (2011). *Acuerdo para la evaluación universal de docentes y directivos en servicio de Educación Básica*. Ciudad de México: SEP/SNTE.
- SNTE. (2008). *Alianza por la calidad de la educación*. Ciudad de México: SNTE.
- SNTE. (2015). *La evaluación del desempeño docente*. Ciudad de México: SNTE.
- Stake, R. E. (2006). *Evaluación Comprensiva y Evaluación Basada en Estándares*. Barcelona: Graó.

Valdés Veloz, H. (2001). *la evaluación del desempeño del docente: un pilar del sistema de evaluación de la calidad de la educación en Cuba*. Recuperado de [http://www.opech.cl/bibliografico/calidad\\_equidad/Valdes%20-%20Evaluacion%20del%20desempeno%20del%20docente.pdf](http://www.opech.cl/bibliografico/calidad_equidad/Valdes%20-%20Evaluacion%20del%20desempeno%20del%20docente.pdf)

Valdés Veloz, H. (2006a). *Desempeño del maestro y su evaluación*. La Habana: Ediciones Pueblo y Educación.

Valdés Veloz, H. (2006b, julio). Los docentes en el sistema educativo cubano: análisis de su carrera, desarrollo profesional y evaluación de su desempeño. Ponencia presentada en el *Encuentro Internacional sobre Evaluación, Carrera y Desarrollo Profesional Docente*, Universidad Cardenal Silva Henríquez de Chile, Santiago de Chile.

Vidal, R. (2009). *¿Enlace, Exani, Excale o PISA?*. Ciudad de México: CENEVAL.

## Breve CV del autor

### Francisco Guzmán Marín

Profesor-Investigador de la Universidad Pedagógica Nacional (UPN). Doctor en Ciencias Sociales, Área de Concentración: Cultura Política y Relaciones de Poder por la UAM-X. Posee una maestría en Filosofía de la Cultura por la UMSNH. Es Licenciado en Matemáticas por la UAT y Profesor de Educación Primaria por la ENR “Vasco de Quiroga”. Ha trabajado 15 años como profesor de Educación Primaria y 18 años como profesor de Educación Superior. ORCID ID: 0000-0001-9433-7680. Email: coraxthelastone@hotmail.com

# Una Propuesta de Ejes para el Análisis del Uso de Resultados de Evaluaciones a Gran Escala del Aprendizaje por Docentes y Directores

## A Proposal of Axes for the Analysis of the Use of Results of Large-Scale Evaluations of Learning by Teachers and Principals

Guadalupe Ruiz Cuéllar \*  
Paloma Prado Robledo

Universidad Autónoma de Aguascalientes - México

Las evaluaciones a gran escala forman parte de los esfuerzos de los sistemas educativos por identificar su situación actual, y, según el caso, aprovechar sus resultados para fines de rendición de cuentas o de mejora pedagógica. Para generar en efecto algún impacto tendiente a la mejora, se requieren procesos de difusión y uso de resultados que aseguren el aprovechamiento de esta información. Este artículo presenta una síntesis de la literatura revisada en el marco de un estudio realizado en Aguascalientes, México, durante 2013, que tuvo como objetivo general, describir los usos que supervisores, directores y docentes hacían de la información derivada de la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE), vigente en el sistema educativo mexicano precisamente hasta ese año. A partir de esa revisión y habiéndolos utilizado en ese estudio, se proponen cinco ejes que pueden orientar la obtención y análisis de información en investigaciones futuras sobre el tema: difusión de la información producida por las evaluaciones a gran escala; análisis e interpretación de la información; uso de los resultados en tres niveles: prácticas docentes, gestión escolar y supervisión; efectos del uso de resultados y factores asociados al aprovechamiento de la información.

**Palabras clave:** Evaluación educativa; Evaluación del aprendizaje; Evaluaciones estandarizadas; Programa de evaluación; Educación básica; Calidad de la educación.

Large-scale assessments are one of the efforts to identify the current situation of the educational systems, and, depending on the case, build on their results for accountability purposes or pedagogical improvement. In order to promote improvement, there is a need to develop and implement specific processes for communicating and using large-scale assessments results. This paper is aimed at presenting a synthesis of the literature reviewed for a study undertaken in Aguascalientes, Mexico, during 2013. The main objective of the study was to describe how the information derived from the Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) was used by supervisors, headteachers and teachers. ENLACE was a census large-scale in force in the Mexican educational system until 2013. As a result of this literature review, five axes are proposed for guiding the collection and analysis of information in future research: dissemination of information produced by large-scale evaluations; analysis and interpretation of information; use of results in teaching practices, school management and supervision; effects of the use of large-scale assessments results; and, factors associated with the use of information.

**Keywords:** Educational assessment; Learning assessment; Standardized assessment; Program evaluation; Basic education; Quality education.

---

\*Contacto: [gruiz@correo.uaa.mx](mailto:gruiz@correo.uaa.mx)

## 1. Introducción

En la actualidad, las evaluaciones a gran escala forman parte de las políticas educativas de un gran número de países, ya sea a través de operativos nacionales o de su participación en proyectos internacionales. Su importancia, entre otros aspectos, se cifra en su capacidad para hacer visibles aspectos fundamentales de los sistemas educativos, tales como el grado en que los alumnos aprenden lo que se espera que adquieran en cada grado o nivel; la medida en que el logro de los aprendizajes se produce de forma equitativa o inequitativa; la diversidad de prácticas educativas entre escuelas y profesores; o el acceso de los estudiantes al conocimiento y los aprendizajes, independientemente de los años de estudio que tengan. Asimismo, aportan información a los diversos actores educativos, misma que puede contribuir a la toma de decisiones y a definir estándares, niveles de desempeño o metas de aprendizaje (Ravela et al., 2007).

Si bien es cierto que estas evaluaciones son relevantes, debido a su capacidad para dar cuenta del estado de los sistemas educativos, también lo es que su aplicación no es suficiente para propiciar una mejora de la realidad que reflejan.

La evaluación no es, en sí misma, un fin. Emplear instrumentos objetivos de medición del aprendizaje que alcanzan los alumnos de una institución o sistema educativo tendrá valor en la medida en que los resultados se aprovechen para el mejoramiento de tales instituciones y sistemas (Martínez Rizo, 2001, p. 83).

Así pues, es necesario que la información derivada de las evaluaciones a gran escala sea difundida de tal manera que llegue a todos los actores educativos para los que está destinada o para quienes sería relevante disponer de la información, y que éstos participen en su análisis e interpretación y hagan un uso adecuado de ella, en aras de propiciar una mejora en las distintas esferas de la educación.

Hace casi dos décadas, Ravela, Wolfe, Valverde y Esquivel (2000) señalaban que una vez instalados formalmente los sistemas nacionales de evaluación en la región de América Latina, aún quedaba mucho por hacer en este ámbito. De manera particular, estos autores indicaban que, en materia de uso y difusión de los resultados de las evaluaciones, se requería “propiciar el desarrollo de trabajos de investigación que permitan recoger evidencia empírica acerca de los efectos que los distintos tipos de reportes de resultados de las evaluaciones nacionales tienen en diversos públicos” (p. 70). Asimismo, enfatizaban la importancia de recabar información sobre lo que los diversos actores educativos esperan de estos sistemas de evaluación, sus percepciones y comprensiones de la información que deriva de éstos, el grado en que la usan, así como sus expectativas y prejuicios acerca de los mismos.

En México, la evaluación externa de resultados de aprendizaje tuvo un gran auge en el periodo 2006-2013, a través de dos programas: la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) y los Exámenes de la Calidad y el Logro Educativo (Excale), aplicados por la Secretaría de Educación Pública (SEP) y el Instituto Nacional para la Evaluación de la Educación (INEE), respectivamente (Backoff y Contreras, 2014; Martínez Rizo, 2015; Ruiz y Prado, 2017). Durante la presente administración federal (2012-2018), el énfasis de la política de evaluación se movió claramente de la evaluación de alumnos a la de docentes en el marco de la implementación en el país del Servicio Profesional Docente, si bien, la primera no dejó de realizarse, pero

a través de un nuevo instrumento, el Plan Nacional para la Evaluación de Aprendizajes (PLANEA), que busca superar las limitaciones de los programas de evaluación que le precedieron, sobre todo ENLACE.

Este artículo presenta una síntesis de la literatura revisada en el marco de un estudio realizado en Aguascalientes, México, durante 2013 -último año en que se aplicó ENLACE-, que tuvo como objetivo general, describir los usos que supervisores, directores y docentes hacían de la información derivada de esta evaluación, única que la producía de manera censal y que, por ende, llegaba a todas las escuelas participantes cada ciclo escolar.

A fin de mostrar de manera sistemática las contribuciones de la literatura científica que guiaron el análisis de la información obtenida en ese estudio, el artículo se divide en dos grandes apartados. En el primero se presentan algunas investigaciones acerca de la difusión y uso de resultados de evaluaciones a gran escala, desarrolladas tanto en América Latina como en México. En el segundo se proponen, incorporando investigaciones adicionales, cinco ejes para el estudio de la difusión y uso de los resultados de las evaluaciones a gran escala: 1) difusión de los resultados; 2) análisis e interpretación de los resultados; 3) uso de los resultados; 4) impacto del uso de los resultados y 5) factores asociados al aprovechamiento de la información de evaluaciones a gran escala. Este último eje es relevante, debido a que su indagación permitiría trascender la mera descripción de los usos y avanzar hacia la comprensión de las razones o factores que condicionan su aprovechamiento en la práctica.

A riesgo de reiterar, pero con el propósito de delimitar los alcances de esta contribución, es importante señalar que los ejes utilizados en el estudio sobre ENLACE, están sustentados en la revisión de los referentes teóricos y conceptuales, así como de los hallazgos de varios estudios previos, realizados tanto en México como en otros países, sobre la difusión y uso de resultados de evaluaciones externas. La pretensión del artículo no es, en sentido estricto, proponer un marco analítico para nuevas investigaciones sobre este objeto de estudio, pero sí aportar algunos elementos que parece relevante considerar por su capacidad informativa del tema, particularmente si el interés está puesto en estudiar los usos que pueden hacer de esos resultados profesores y directores, y si las pruebas a que corresponden dan prioridad al propósito de “brindar información oportuna y relevante a los docentes que les ayude a mejorar tanto su enseñanza como el aprendizaje de sus estudiantes (función pedagógica)” (Backhoff y Contreras, 2014, p. 1269).

Finalmente, cabe añadir que en lo que resta del texto se hace alusión de manera indistinta a los términos “evaluación” y “prueba”, pero se reconoce que no son equivalentes. Se coincide en este sentido con el argumento de Moreno (2011), quien precisa a propósito de un texto suyo:

*Este trabajo se alude a las pruebas estandarizadas como evaluación porque así es como habitualmente se les denomina, pero es preciso advertir que no son sinónimos; las pruebas o exámenes son uno más de los diversos instrumentos que la evaluación puede emplear para la valoración de un determinado objeto. (p. 117)*

## 2. Estudios previos

En América Latina se han realizado diversas investigaciones orientadas a dar cuenta de algunos aspectos relativos al uso de resultados de las evaluaciones a gran escala. Una de estas iniciativas fue llevada a cabo en 1999 por el Grupo de Trabajo sobre Estándares y

Evaluación (GTEE) del PREAL, cuyo objetivo fue analizar la situación de los sistemas nacionales de evaluación de diez países de la región: Argentina, Brasil, Bolivia, Costa Rica, Chile, Ecuador, México, Paraguay, Uruguay y Perú.

Fruto de este proyecto es el documento *¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?* (Ravela, 2001). Uno de los temas abordados, referido a la información producida y difundida por los sistemas nacionales de evaluación, es el de los usos que se pretende se hagan de este tipo de datos y si efectivamente se llevan a la práctica.

Los hallazgos de esta investigación muestran que la mayor parte de los sistemas nacionales de evaluación de los aprendizajes en América Latina son de carácter formativo y su uso es de bajo riesgo,<sup>1</sup> salvo algunas excepciones como Chile y México,<sup>2</sup> que implementaron sistemas cuyos usos eran de alto riesgo.<sup>3</sup> Ambos tipos tienen desventajas, dado que los usos de consecuencias fuertes o alto riesgo pueden generar que los actores educativos dediquen una cantidad excesiva de tiempo a la preparación para las pruebas, reduciendo el currículo (Backhoff y Contreras, 2014, Ravela, 2001); mientras que los usos de bajo riesgo o la ausencia de consecuencias a partir de los resultados de las evaluaciones pueden propiciar poco interés y responsabilidad por los resultados de este tipo de pruebas. En las conclusiones de este trabajo el autor menciona, entre otros aspectos, que es importante realizar investigaciones acerca de los efectos que tiene el uso de las evaluaciones a gran escala para establecer incentivos.

El Programa de Promoción de la Reforma Educativa en América Latina y el Caribe (PREAL), con el apoyo del Banco Mundial abrió, en 2002, la primera convocatoria del Fondo de Investigaciones Educativas (FIE). El propósito fue financiar proyectos en torno a la siguiente pregunta: “¿Cómo se usa y qué impacto tiene la información empírica en el mejoramiento de los sistemas educativos en América Latina?” (Cueto, 2005, pp. 11-12). Los resultados de algunos de los estudios realizados se sintetizan enseguida.

Fernández y Midaglia (2005) analizaron el uso de informes de evaluación en dos países cuyos sistemas de evaluación tiene características diferentes: México y Uruguay. Para ello, se tomaron en cuenta los aspectos organizacionales de cada centro educativo y se generó una tipología de la gestión de la información. Fernández y Midaglia (2005) concluyeron

---

<sup>1</sup> Este tipo de evaluaciones tiene un carácter formativo y un uso de bajo impacto, esto se refiere a que se emplean principalmente para dar información acerca de los sistemas educativos, ayudar a la elaboración de políticas públicas y propiciar la mejora de las instituciones escolares, así como el desarrollo profesional de los docentes (Ravela, 2001).

<sup>2</sup> La investigación considera el caso del Programa “Carrera Magisterial” y la asignación que conllevaba de incentivos económicos para docentes, considerando, entre otros factores, los resultados de sus alumnos en la evaluación ENLACE. El sistema de evaluación que actualmente se aplica en México (PLANEA) no tiene asociado este uso.

<sup>3</sup> Ravela (2001) señala que a diferencia de las evaluaciones de bajo riesgo, las de alto impacto tienen un carácter sumativo y se emplean principalmente para la acreditación y selección de alumnos, el establecimiento de incentivos para docentes e instituciones escolares y la creación de *rankings* de escuelas.

que en ambos países la lectura de informes de evaluación es escasa y, de manera general, los docentes que leen los informes son los que exhiben un mayor nivel de profesionalismo. La principal dificultad reportada por los docentes para leer los informes es la falta de orientaciones claras del director del centro o del supervisor.

Aristimuño y Kaztman (2005), a través de entrevistas en profundidad y grupos focales, indagaron en torno al uso y comprensión de los informes de evaluación del rendimiento de los estudiantes en Uruguay. Los autores describen las reformas educativas emprendidas en el país, especialmente a partir de 1995. Dos características importantes del sistema uruguayo consisten en que los informes de resultados son utilizados para impactar directamente en el trabajo de los maestros y directores, pues son retomados en cursos de formación en servicio y talleres, y en reuniones que tienen lugar en las mismas escuelas. Por otra parte, las escuelas que no han participado en alguna evaluación nacional reciben un informe de los resultados, así como la prueba para que ésta sea autoadministrada en las instituciones y comparada con varios referentes (por ejemplo, el promedio nacional o el promedio de las escuelas del mismo contexto sociocultural). Estas prácticas, según los autores, brindan legitimidad al operativo de evaluación entre docentes y directivos, sin embargo, encuentran también un uso insuficiente de los informes de evaluación. A partir de sus hallazgos, Aristimuño y Kaztman (2005) formulan un modelo conceptual que tiene como variable a explicar la apertura al cambio del maestro (la disposición para usar los informes de evaluación), y como variables explicativas (relacionadas entre sí) la antigüedad del docente, el tipo de escuela en que trabaja y la densidad de las redes sociales de la escuela.

Montoya, Perusia y Vera (2005) analizaron el uso de la información sobre las evaluaciones del rendimiento escolar realizadas por el Ministerio de Educación Nacional en Argentina. Los autores encontraron un conocimiento diferencial de los informes según la asignatura a la que se refieren (mayor en el caso de español y matemáticas y menor en el de otras asignaturas evaluadas). La oportunidad con que llegó la información a los establecimientos también fue diferente según las pautas de entrega seguidas en tres provincias distintas. Se advierte que no existe un plan que guíe a las instituciones respecto al uso de la información proporcionada. Los autores concluyen que para que los informes de evaluación sean insumos relevantes para el cambio debe trabajarse con los docentes en la interpretación de la información reportada y en el análisis de alternativas pertinentes a los resultados.

Aunque estos trabajos abordaron desde diferentes ángulos un mismo objeto de estudio y emplearon para ellos diversas metodologías, Cueto (2005) señala que, en conjunto, estas investigaciones dejan ver que en América Latina prevalece una preocupación por la información que se debe recabar y la manera de hacerlo, mientras que hay poco desarrollo a propósito de cómo emplearla para contribuir a la mejora de la educación. De acuerdo con el autor, pese a que estos sistemas fungen como insumos de información que es vital para los sistemas educativos, hay poco uso de estos datos por parte de las audiencias a las que se dirigen. Para Martinic (2010) esto deriva, en parte, de “las insuficiencias en la comunicación hacia los actores y [la falta de] interpretaciones que les ayuden a comprender los problemas que verifican los datos reunidos” (p. 38).

Finalmente, cabe hacer alusión al informe “Las evaluaciones educativas que América Latina necesita” (Ravela, 2010), documento que derivó del proyecto “La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas”, llevado a

cabo por el Instituto de Evaluación Educativa de la Universidad Católica del Uruguay, con apoyo del Grupo de Trabajo sobre Estándares y Evaluación (GTEE) de PREAL. Este trabajo fue un estudio exploratorio y descriptivo elaborado en ocho países de América Latina (Argentina, Colombia, Costa Rica, El Salvador, México, Perú y Uruguay) con maestros de sexto grado de primaria, cuyo énfasis estuvo en los enfoques de la evaluación en las áreas de matemáticas, lectura y escritura.

El trabajo tuvo entre sus objetivos estudiar cómo las evaluaciones a gran escala ayudan a los profesores a enriquecer sus propias prácticas de evaluación. Gran parte de los docentes indicaron que los resultados de este tipo de evaluaciones efectivamente contribuyen a la mejora de esas prácticas. A partir de la información reunida, se generó una tipología de usos, misma que incluye la toma de conciencia acerca de los contenidos a abordar en el aula; la apertura a nuevos enfoques conceptuales y prácticos; el uso de instrumentos para implementarlos en el aula; la creación de nuevos instrumentos de evaluación; la preparación de los alumnos para las mismas pruebas en aplicaciones posteriores; y la creación de procesos de seguimiento para los alumnos, de manera general o individual (Ravela, 2010).

De igual manera, el estudio identificó factores que pueden afectar el aprovechamiento que los profesores hacen de la información derivada de evaluaciones a gran escala, tales como su comprensión de los datos, la formación de los propios docentes, el exceso de carga de trabajo, las fallas técnicas de las pruebas, las opiniones de los profesores acerca de las evaluaciones, entre otros (Ravela, 2010).

En el contexto mexicano son pocas las investigaciones que se han dado a la tarea de indagar el uso de resultados de las evaluaciones a gran escala, si bien, este campo de estudio ha cobrado mayor importancia en las últimas décadas. Uno de los trabajos que abordó este tema es previo a la implementación de ENLACE y su objetivo fue “[...] captar, registrar y sistematizar significados que profesores, directores, supervisores y funcionarios otorgan a la evaluación externa a partir de su experiencia” (Figueroa, 2003, p. 354). Los resultados de esta investigación apuntan a que existían “pocas y aisladas prácticas de difusión y uso de los resultados de los programas” (p. 372).

Al cabo de algunos años del surgimiento de ENLACE, empezaron a realizarse estudios sobre diferentes aspectos y dimensiones de la evaluación y de la prueba en particular (Martínez Rizo y otros, 2015; Santiago et al., 2014;). A propósito del uso de resultados, Santiago y otros (2014) señalan. en los hallazgos de su investigación, que en sus orígenes ENLACE fue planteada como una evaluación de carácter formativo, cuyos usos eran de bajo impacto; sin embargo, debido a los usos reales que se hicieron de esta información, entre ellos la asignación de incentivos para docentes y la difusión pública de los resultados de las escuelas, se convirtió en una evaluación de alto impacto. La preparación para la prueba, o la reducción del currículo, prácticas antes mencionadas, se hicieron patentes en el caso de ENLACE. Por otra parte, esta prueba también propició una toma de conciencia por parte de los profesores acerca de la necesidad de fortalecer la preparación de los alumnos en determinadas áreas y sirvió como motivación para los estudiantes y los docentes.

Martínez Rizo (2015), con la colaboración de un grupo internacional de expertos coordinó un estudio de validación de las pruebas ENLACE para educación básica (ENLACE-B) y media superior (ENLACE-MS), así como de los Exámenes para la Calidad y el Logro

Educativo (Excale). Una de las dimensiones incluidas en el estudio es la del uso de los resultados; entre los hallazgos principales cabe mencionar los siguientes: la carencia de un modelo que pueda alinear los aspectos técnicos de las pruebas a sus usos previstos, así como la diversidad de usos y su nivel de especificidad; la falta de acceso a la información de manera oportuna y efectiva; la carencia de organización y coordinación para fomentar la capacidad de los diversos actores educativos en la interpretación de los datos; el poco desarrollo de mecanismos eficientes y robustos que permitan a los investigadores acceder y hacer uso de estas bases de datos; y la falta de seguimiento a los usos y consecuencias derivadas de éstos.

### **3. Ejes para el estudio de la difusión y uso de resultados de evaluaciones a gran escala**

Los ejes que se presentan en esta parte del artículo fueron identificados en la revisión de literatura realizada en el marco del proyecto de investigación al que se hace referencia en la Introducción. Estos ejes fueron de utilidad para sistematizar la información reunida y se considera que resultan sugerentes para indagaciones posteriores sobre el tema, advirtiendo que quizá sería necesario revisar si los ejes propuestos son suficientes para abarcar de forma más exhaustiva el tema y si son pertinentes dada la evaluación que se esté considerando en particular. Aunque el proyecto en que fueron utilizados como referentes teóricos centró su atención en los usos de los resultados por profesores y directores de escuelas primarias, se incluyen dos ejes previos a esta acción, la difusión y el análisis e interpretación de los datos, debido a que, en la literatura revisada, aparecen como procesos de claro impacto sobre el uso que efectivamente se puede hacer de los resultados; igualmente se plantea un eje relacionado con los factores que influyen en su uso por parte de los actores mencionados.

#### ***3.1. Difusión de la información de evaluaciones a gran escala***

La evaluación cobra sentido en la medida en que contribuye a que la educación mejore (INEE, 2006b) y para lograr esto, entre otras acciones, es necesario que los resultados se difundan de manera oportuna y amplia a todos los actores educativos, quienes, a su vez, deben poseer una adecuada cultura de la evaluación, entendiendo este término como:

*[...] Un conocimiento de las características, alcances y limitaciones de los distintos tipos de evaluación, con base en lo cual tanto las autoridades como los maestros, los padres de familia y la sociedad en general puedan interpretar los resultados de las evaluaciones correctamente, y los puedan utilizar para sustentar decisiones adecuadas, en los respectivos ámbitos de cada uno. (Martínez Rizo y Santos, 2009, p. 10)*

La difusión de resultados depende del enfoque y las concepciones asumidas por los sistemas de evaluación nacional y los fines para los que fueron creados. Ravela (2012) indica que hay dos enfoques para llevar a cabo esta tarea: data-céntrico y agere-céntrico. En el primero de éstos el punto central son los datos de la evaluación, vistos desde la lógica de la estadística y la psicometría; se procesan los datos obtenidos y se generan informes. Este enfoque enfatiza la difusión de los resultados, pero no su uso, dado que éste se encuentra en manos de otros actores. Desde esta postura, se asume que la información estadística compleja podrá ser comprendida por todos los actores educativos y que sabrán como emplearla, además de presuponer que la información será de utilidad para la toma

de decisiones y que los usuarios tienen tanto los incentivos como la motivación para usar la información.

A diferencia del anterior, el segundo enfoque se orienta a los actores, de tal manera que se busca, en primera instancia, familiarizarse con sus modos de actuar respecto a la educación, sus percepciones e intereses, así como con el tipo de decisiones que pueden tomar. Con base en esto, se decide cómo se hará el tratamiento de los datos, buscando difundir la información que les pueda ser útil a los actores para tomar decisiones. El enfoque agerecéntrico es menos usual en América Latina, dado que los sistemas de evaluación nacionales han puesto mayor énfasis en la difusión de los resultados que en su uso (Ravela, 2012).

A partir de un estudio acerca del uso de la información de los sistemas de evaluación de educación primaria en México y Uruguay, Fernández y Midaglia (2005) también señalan que hay diversas estrategias para comunicar los resultados. La más frecuente se denomina *trickle down* o de chorreo y consiste fundamentalmente en la distribución de los reportes de resultados de uno a otro actor, comenzando por las autoridades educativas o instancias encargadas de iniciar esta tarea, hasta llegar a las escuelas<sup>4</sup>.

La segunda estrategia es la *resignificación*, que parte del supuesto “[...] de la comunicación transparente y someter en distintas instancias a reflexión el sentido de lo comunicado” (Fernández y Midaglia, 2005, p. 129). Dicha estrategia tiene dos subtipos, en el primero, el informe de los resultados se acompaña por un documento de trabajo dirigido a profesores y directores, el cual puede referirse a los objetivos de la evaluación o a la interpretación de los datos. El segundo subtipo tiene como punto de partida un cuestionamiento, ya sea implícito o explícito, y en este sentido, “[...] las organizaciones intermedias simplemente adaptarían y aprehenderían los informes bajo las premisas estructuralmente disponibles, que son típicamente burocráticas” (p. 130).

El último tipo de difusión de resultados es llamado estrategia de *reestructura*, la cual se define de la siguiente manera:

*[...] una política de reforma estructural de la supervisión escolar que pone en cuestión las relaciones interorganizacionales entre gerencia, dirección de supervisión y escuelas. Entre otros impactos deseados, contribuiría a jerarquizar el uso de las evaluaciones como un insumo de las prácticas de asesoramiento académico, de diagnóstico y monitoreo de las escuelas y sus maestros.* (Fernández y Midaglia, 2005, p. 130)

### **3.2. Análisis e interpretación de la información**

Uno de los aspectos a abordar cuando se indagan los usos que diversos actores hacen de los resultados de pruebas aplicadas en evaluaciones a gran escala es el análisis que realizan a partir de la información que reciben, y lo que este proceso les permite ver, es decir, las interpretaciones de los resultados, mismas que influirán previsiblemente, en la toma de decisiones.

Cawelti y Protheroe (2000, en Protheroe, 2009), Fernández y Midaglia (2005), Means y otros (2009), Ravela (2010) y Ravela y otros (2007) señalan, como parte de los usos de los

---

<sup>4</sup> La investigación realizada por Fernández y Midaglia tomó como referente en el caso de México, la evaluación Aprovechamiento Escolar, que formaba parte del Programa Carrera Magisterial.

resultados de las pruebas a gran escala, la generación de diagnósticos del aprendizaje de los estudiantes y/o de los grupos, así como del funcionamiento de las escuelas o regiones.

En el estudio llevado a cabo por el Council of the Great City Schools (2012), se señala que dar sentido a los datos provenientes de las evaluaciones intermedias y finales aplicadas a los alumnos desemboca idealmente, en una mejora del conocimiento y comprensión de las necesidades y retos de los alumnos, de manera individual y como grupo; de las fortalezas y áreas de oportunidad de los profesores y las escuelas, así como de las estrategias y recursos para dar atención a las dificultades de los alumnos, los profesores y las escuelas (Council of the Great City Schools, 2012).

A propósito de esta misma cuestión, Griffin (2009) indica que las evaluaciones cuyos fines son formativos pueden contribuir a identificar el nivel de desarrollo de los estudiantes y, con base en esta información, es posible implementar intervenciones que contribuyan al progreso de los alumnos. Para llevar a cabo esto se requiere, como condiciones indispensables, que el propio instrumento de evaluación posea suficiente capacidad para generar un diagnóstico del aprendizaje de los alumnos y que los profesores y, de manera general, los actores educativos que reciben, analizan y usan los resultados, tengan habilidades para la interpretación y el uso de este tipo de datos (Griffin, 2009).

Matters (2009) indica que las pruebas estandarizadas pueden ofrecer información a propósito no sólo de lo que los alumnos aprendieron, sino también de aquello que no aprendieron. Ante esto, es necesario que los docentes puedan identificar las condiciones en las cuales los alumnos llegaron o no a determinado nivel de logro. Love (2004) también indica que en el proceso de uso de datos se deben identificar los problemas de los estudiantes y verificar sus causas.

Al igual que Love (2004), el Institute of Education Sciences (2009) señala que la interpretación de la información derivada de este tipo de evaluaciones consiste en identificar las fortalezas y debilidades de los alumnos, ya sea de forma individual o del grupo, y, subsecuentemente, desarrollar hipótesis acerca de cómo mejorar el aprendizaje de los alumnos.

Tanto el Institute of Education Sciences (2009) como Love (2004) y Ravela (2001) señalan la importancia de triangular información acerca del aprendizaje de los alumnos, con el fin de comprobar los datos que ofrecen distintas fuentes y así llegar a conclusiones fundamentadas acerca de las necesidades educativas de los alumnos.

### ***3.3. Uso de los resultados***

Si bien es cierto que la difusión de los resultados de las evaluaciones a gran escala o estandarizadas y su análisis e interpretación son condiciones indispensables para asegurar su adecuado aprovechamiento, el uso que se haga de esta información es lo que da sentido a la propia tarea de evaluar.

Además de la distinción ya mencionada entre usos de alto y bajo riesgo o impacto, a partir de la literatura consultada también se pueden distinguir tres niveles en los que se usan los resultados de las evaluaciones a gran escala y diversos matices u orientaciones en cada caso.

### 3.3.1. Nivel de las prácticas docentes

En el primer nivel se engloban todos aquellos usos y acciones que se espera que los profesores lleven a cabo desde sus prácticas docentes, ya sean de enseñanza o evaluación. Dentro de este nivel, se encuentran diversos planteamientos que se pueden agrupar en cinco categorías:

1. Planeación de clase
2. Modificación y mejora e intervenciones
3. Prácticas de evaluación
4. Apoyo y seguimiento de estudiantes
5. Pruebas estandarizadas

El diseño y/o ajuste de la planeación de clase incluye diversos aspectos que van desde la modificación de los planes de clase (Council of The Great City Schools, 2012), reordenar los planes de estudios, su secuencia y el tiempo dedicado a cada tema (Institute of Education Sciences, 2009, Hoover, 2009, en Council of The Great City Schools, 2012), hasta cuidar la alineación del currículo o de las expectativas de aprendizaje (Institute of Education Sciences, 2009, Protheroe, 2009).

La segunda categoría está sustentada en las contribuciones de autores como Coburn y Talbert (2006), Hoover (2009, en Council of the Great City Schools, 2012) y Protheroe (2009), que se refieren a la modificación, mejora o diferenciación de las prácticas y enfoques de enseñanza de los profesores. Las intervenciones consisten en el diseño e implementación de diversas estrategias y actividades. En algunos casos se habla de manera general de la realización de “intervenciones” (Griffin, 2009; Institute of Education Sciences, 2009), mientras que, en otros, se señala de manera específica, que un uso que se da a la información de este tipo de evaluaciones consiste en volver a enseñar determinados temas o conceptos (Hoover, 2009).

La tercera categoría alude a la modificación de las prácticas evaluativas de los docentes, ya se trate de retomar los instrumentos de las evaluaciones a gran escala para aplicarlas en el aula o que éstas sirvan como modelo para generar nuevas evaluaciones (Ravela, 2010). Autores como Cawelti y Protheroe (2000, en Protheroe, 2009) señalan que, a partir de los resultados de este tipo de pruebas, se pueden aplicar evaluaciones intermedias; mientras que Hoover (2009, en Council of The Great City Schools, 2012) menciona que una práctica común es aplicar nuevamente la evaluación para abordar determinados temas.

La cuarta categoría comprende el apoyo y seguimiento de los estudiantes con bajos niveles de logro. Diversos autores señalan que, a partir del diagnóstico generado con base en la información de las pruebas, se puede dar seguimiento o tutoría a los alumnos que presentan dificultades, así como ubicarlos en grupos especiales para brindarles apoyo adicional (Cawelti y Protheroe, 2000, en Protheroe, 2009; Council of The Great City Schools, 2012; Insitute of Education Sciences, 2009; Hoover, 2009, en *Council of The Great City Schools*, 2012; Ravela, 2010).

Finalmente, Ravela (2010) también incluye, como uno de los usos de los resultados de evaluaciones a gran escala, la preparación para este tipo de pruebas, aunque el mismo autor indica que:

*Este uso puede ser legítimo si la finalidad es familiarizar a los alumnos con un cierto tipo de preguntas y de evaluaciones, pero tiene el problema de que muchas veces termina transformando en el objetivo principal de la enseñanza lo que pretende ser solo un indicado*  
(p. 14)

### 3.3.2. Nivel de la gestión escolar

En el segundo nivel se encuentran todos aquellos usos de los resultados de evaluaciones a gran escala que son atribuidos al colectivo escolar. Autores como Ravela (2001) señalan, de manera general, que los resultados de evaluaciones a gran escala pueden contribuir al mejoramiento de las escuelas, aunque no se especifica a través de qué medios se lleva a cabo esto. Algunas otras contribuciones llevan a identificar las siguientes categorías en este nivel:

1. Mejora escolar, planes escolares e intervenciones a nivel escuela
2. Revisión del currículo
3. Comunicación con los padres

En la primera categoría se encuentran las contribuciones de Ravela (2001) y de Santiago y otros (2014), autores que señalan que los resultados deben usarse para propiciar el fortalecimiento o mejora de las escuelas. También comprende el uso de resultados para elaborar o ajustar planes escolares o de mejora, así como para plantear nuevas metas (Institute of Education Sciences, 2009; Department of Education, 2010, en Council of The Great City Schools, 2012 y Council of The Great City Schools, 2012). Finalmente, en esta categoría se pueden ubicar usos relacionados con la implementación de estrategias y actividades a nivel escolar (Henke, 2005, en Council of The Great City Schools, 2012 y Council of The Great City Schools, 2012).

La segunda categoría alude a la revisión y ajuste del currículo a partir de la información de pruebas a gran escala (Department of Education, 2010, en Council of The Great City Schools, 2012 y Council of The Great City Schools, 2012), lo cual tiene sentido en sistemas educativos como el norteamericano donde el currículum se encuentra descentralizado.

Finalmente, la tercera categoría consiste en informar a los padres de los alumnos acerca del progreso académico de sus hijos (Means et al., 2009). Este punto forma parte de los propósitos y usos sugeridos de ENLACE, que de manera particular proponía la realización de actividades, tales como talleres para propiciar la participación de los padres en el proceso de aprendizaje de los alumnos.

### 3.3.3. Nivel de la supervisión

En el nivel de la supervisión se incluyen aquellos usos de la información que hacen las instancias externas a las escuelas; en este caso también se advierten diferentes posibilidades:

1. Desarrollo profesional docente
2. Apoyo a las instituciones escolares
3. Diseño de políticas públicas

La primera categoría engloba, de manera general, los usos encaminados a fortalecer el desarrollo de los docentes a través de diversas estrategias. En la investigación de Council of The Great City Schools (2012), se habla de las acciones que emprenden los distritos con

base en los resultados de evaluaciones intermedias y finales del aprendizaje de los alumnos y entre ellas se encuentran las estrategias encaminadas al desarrollo profesional de los docentes. En este punto coinciden Ravela (2001), SEP (2013) y Santiago y otros (2014), al igual que Kerr, Marsh, Darilek y Barney (2006, en Young y Kim, 2010).

La segunda categoría se refiere al apoyo a las instituciones escolares, que es uno de los usos a los que hace alusión la investigación de Council of The Great City Schools (2012), particularmente el acompañamiento a las escuelas en la generación del plan de mejora escolar. Por último, la tercera categoría comprende la generación de programas, políticas públicas y educativas por parte de las autoridades escolares o del gobierno (Coburn y Talbert, 2006; Dussel et al., 2005; Ravela, 2001; Santiago et al., 2014).

### **3.4. Efectos del uso de los resultados**

Los usos que se den a los resultados de las evaluaciones a gran escala producen un efecto o impacto en diferentes niveles de los sistemas educativos. Con respecto a este tema, los modelos de uso de resultados presentados por Love (2004) y el Institute of Education Sciences (2009) señalan que el efecto esperado del proceso de analizar, interpretar y usar los datos de este tipo de evaluaciones consiste en una mejora tanto en la enseñanza como en el aprendizaje de los alumnos.

Para Ravela (2010), el uso de los resultados de pruebas estandarizadas puede propiciar que los profesores tomen conciencia de los conocimientos y capacidades que los alumnos deben adquirir. Por otra parte, contribuye a que los docentes se abran ante otros horizontes conceptuales en el ámbito de la didáctica y la evaluación; les brinda acceso a instrumentos de evaluación que les pueden servir de apoyo en el trabajo en el aula o para crear nuevos instrumentos; y les permite reflexionar a propósito de sus prácticas de evaluación en el aula, poniendo mayor énfasis en los procesos cognitivos de sus alumnos.

Si bien lo deseable es que el uso de este tipo de evaluaciones derive en una mejora general de los sistemas educativos, hay usos que desembocan en efectos no deseados. Matters (2009) indica que una acción común es que las ciudades y países se definan en función de los resultados de evaluaciones como PISA.

Por su parte, Dussel y otros (2005) encontraron, a partir de una investigación realizada en las provincias de Córdoba, Chaco y la Ciudad Autónoma de Buenos Aires en Argentina, que los docentes consideran que el impacto de los sistemas de evaluación ha sido poco visible o incluso nulo, debido a situaciones tales como “[...] la falta de acceso, el desconocimiento, la escasa necesidad de estadísticas para la labor docente, la falta de presupuesto para modificar situaciones ya identificadas como problemáticas, o porque cada docente está en sus temas específicos” (p. 377).

Matters (2009) afirma que, en este siglo, la información derivada de las evaluaciones es extremadamente poderosa, siempre y cuando cumpla con dos condiciones. La primera de ellas consiste en que la prueba, por sí misma, sea de calidad y lo suficientemente fuerte para fundamentar los usos que se hacen a partir de la información que arroja; mientras que la segunda condición señala que los usuarios de esta información deben poseer suficiente pericia para reconstruir adecuadamente la realidad que subyace a los resultados.

### ***3.5. Factores asociados al aprovechamiento de la información de evaluaciones a gran escala***

La mayor parte de la literatura revisada en esta investigación refiere una amplia cantidad de factores que impactan en el uso de los resultados de evaluaciones a gran escala, mismos que se pueden agrupar en dos categorías los factores de orden individual, relativos a los actores educativos, y los que se ubican en el nivel de la institución escolar.

#### *3.5.1. Factores individuales*

Los factores individuales se definen como las percepciones y actitudes de los actores educativos, mismas que pueden favorecer u obstaculizar el trabajo con los datos. Cabe señalar que esto no alude ni involucra elementos relacionados con el contexto sociodemográfico de los docentes, variables que diversas investigaciones (Fernández y Midaglia, 2005) han demostrado no intervenir en aspectos como la lectura de los resultados por parte de los profesores.

Ravela (2010) afirma que es pertinente considerar las percepciones de los docentes, dado que “[...] la evaluación externa puede ser percibida de distintas maneras: como una amenaza, como una oportunidad de aprendizaje o como un requisito administrativo más [...]” (p. 18). Esto depende de la cultura de evaluación de cada institución y de la estrategia de comunicación de los resultados, pero también se atribuye a las ideas previas de los docentes. Asimismo, factores como la formación inicial de los profesores, el clima institucional y profesional de las escuelas, la complejidad de la labor docente “[...] pueden favorecer u obstaculizar la comprensión y el uso de evaluaciones [...]” deben ser tenidas en cuenta a la hora de promover el uso de los resultados de las evaluaciones” (Ravela, 2010, p. 18). Datnow, Park y Wohlstetter (2007, en Young y Kim, 2010) enfatizan que la capacidad de los docentes para usar los datos de las evaluaciones está “inextricablemente ligado a su conocimiento instruccional. El análisis de los datos puede indicar las áreas en las que los profesores deben hacer mayores esfuerzos, pero no les dicen cómo hacerlo” (p. 10).

En la investigación de Council of The Great City Schools (2012) también se indica que las percepciones negativas a propósito de la calidad de los datos de la evaluación pueden llevar a que los actores educativos se muestren renuentes ante las estrategias implementadas a partir de esta información, lo que trae consigo una disminución del aprovechamiento de los resultados de las evaluaciones. De acuerdo con Young y Kim (2010), “aunque los distritos pueden usar muchas estrategias para estimular y apoyar a las escuelas en el uso de los datos, éstos deben tener legitimidad ante los profesores” (p. 18) para que los usen.

La investigación de Council of The Great City Schools (2012) reporta que otro tipo de factores que intervienen en el uso de las evaluaciones intermedias y finales son las limitaciones de las capacidades del personal, mismas que en otros estudios también se ha encontrado que obstaculizan el uso de los resultados (Departamento de Educación de Estados Unidos, 2010a; Herencia, 2007; Herman et al., 2008; Ingram et al., 2004; Lachat y Smith, 2005; Sharkey y Murnane, 2006; Patrimonio y Bailey, 2006; Wayman et al., 2007). En el mismo sentido, esto es, las debilidades en la formación de los profesores para la evaluación, apuntan las conclusiones de Impara y otros (1993) y Stiggings (1991) (citados en Young y Kim, 2010).

Por su parte, Dussel y otros (2005) señalan que hay varios factores relevantes para poder interpretar el comportamiento de los docentes ante la información empírica, uno de éstos se relaciona con el escepticismo de los profesores ante la utilidad de la información. De igual manera, estos autores reportan diversas opiniones ante las estadísticas que reportan las evaluaciones, algunas de ellas positivas y otras que apuntan a una conceptualización de este tipo de información como datos en los que hay sesgo y que, por tanto, no reflejan la complejidad de la realidad evaluada. Finalmente, también hay opiniones que los autores denominan “ausentes” y consiste en una postura que concibe las estadísticas como una actividad carente de sentido.

### 3.5.2. Factores institucionales

Los factores institucionales se definen como aquellos aspectos de la dinámica escolar o de la logística de la evaluación que intervienen en el aprovechamiento de los resultados de evaluaciones a gran escala. La literatura consultada en este estudio apunta principalmente a tres factores que puede determinar la difusión, análisis, interpretación y uso de los resultados: 1) la entrega oportuna de la información, 2) el trabajo colegiado, y 3) la existencia de liderazgos.

A propósito del primer punto, la investigación de Council of the Great City Schools (2012) señala que diversos autores (Clune y White, 2008; Lachar y Smith, 2005; Wayman et al., 2007) han encontrado que algunas de las barreras para el uso de los resultados de las evaluaciones intermedias es la falta de puntualidad y accesibilidad en la entrega de los resultados, así como la carencia de precisión de éstos. Rojas y Esquivel (1988, en Cruz, Montoya y Vera, 2005) coinciden en este punto y afirman que, de no hacer una entrega oportuna de los resultados, éste se torna “[...] un ejercicio mecánico y costoso, que genera una gran cantidad de información que en últimas no es utilizada, hecho que incide en la seriedad y prestigio que se le asigne al proceso” (p. 88).

Sobre el trabajo colegiado, son varios los investigadores (Council of the Great City Schools, 2012; Griffin, 2009; Ravela, 2010) que señalan la importancia de que exista una dinámica de trabajo colectivo en torno al análisis, interpretación y uso de los resultados de evaluación. Con respecto a este tema, Griffin (2009) indica que los equipos profesionales de enseñanza permiten a los profesores exponer y compartir sus interpretaciones acerca de los datos, así como proponer estrategias e intervenciones y, sobre todo, hacer un trabajo reflexivo y crítico en torno a estos aspectos.

El modelo de uso de evaluaciones expuesto por el Institute of Education *Sciences* (2009) también enfatiza la importancia de establecer equipos de análisis de datos que dirijan su uso y definan conceptos cruciales para ello, tales como las nociones de enseñanza y aprendizaje. Por otra parte, en la investigación de Council of the Great City Schools (2012) se indica que una de las barreras para el uso de los datos, de acuerdo con el Department of Education (2010), es la falta de tiempo para participar en la examinación y reflexión de los resultados.

Finalmente, la literatura también apunta a la existencia de liderazgos dentro de las instituciones escolares como un elemento que incide en el proceso del uso de datos de las evaluaciones a gran escala. Fernández y Midaglia (2005) establecieron como una de las hipótesis de su investigación que: “los estilos de gestión del director de la escuela tienen efectos diferenciales sobre la lectura y uso de los informes por parte de los maestros” (p.

157). Los resultados de este estudio indicaron que para muchos de los actores educativos encuestados el éxito de la lectura de los reportes de resultados “[...] no dependería de características personales, sino principalmente del director y de su capacidad de motivación, fomento de la lectura, supervisión de la labor docente, su rol en la formación de equipos estables y académicamente sólidos, entre otros aspectos” (p. 156).

A propósito de este tema, el Institute of Education Sciences (2009) señala que es indispensable proporcionar liderazgo en el proceso de uso de datos. Asimismo, Christman y otros (2009) y Kerr y otros (2006), ambos citados en Council of the Great City Schools (2012) afirman que uno de los factores determinantes en el uso de resultados, en este caso de evaluaciones intermedias, es el liderazgo a nivel de enseñanza y a nivel escolar.

Con base en diversos autores, Young y Kim (2010) establecen que:

*El liderazgo del director es crucial para establecer expectativas entre el colectivo escolar sobre el uso de los datos como insumos para la toma de decisiones (Wayman y Stringfield, 2006a, Young, 2006) y para crear ambientes de apoyo en los que los profesores puedan compartir los éxitos y los fracasos asociados con la evaluación. (p. 19)*

Además de expectativas, un adecuado liderazgo precisa las actividades a desarrollar, modela los usos apropiados de la información, proporciona una teoría de la acción para el uso de los datos y despliega los recursos necesarios para hacer todo esto (Young y Kim, 2010).

#### 4. Conclusiones

Durante los últimos años se han dado notables avances en materia de difusión de resultados de las evaluaciones a gran escala. En 2006, el INEE afirmaba que instancias como la Subsecretaría de Educación Básica (SEB), la Dirección General de Evaluación de Políticas (DGEP) y el propio INEE habían contribuido a ampliar el acceso a la información a través de diversos materiales (publicaciones, análisis de resultados, manuales técnicos, entre otros). A pesar de esto, el INEE (2006b) llamaba la atención sobre la inexistencia de mecanismos sistemáticos de difusión entre los actores educativos: alumnos y sus familias, maestros y directores:

*Lógicamente, esto impide que los resultados sean utilizados en forma sistemática para mejorar los procesos educativos básicos, es decir las actividades de aprendizaje. A esto se agrega que, cuando los resultados son difundidos hacia la sociedad a través de los medios de comunicación, suelen generarse debates estériles derivados de interpretaciones incorrectas o sesgadas. (p. 2)*

En el marco de la evaluación ENLACE, la difusión de los resultados se constituyó como elemento fundamental, lo que contribuyó a su gran notoriedad entre la sociedad en general y en el sector educativo, a pesar de que muchas de las dificultades en el aprovechamiento de esta información se originaban precisamente desde esta etapa. A propósito de los mecanismos de difusión de resultados de ENLACE y Excale, Martínez Rizo (2015) señala que “[...] se encontraron carencias técnicas e inconsistencias entre la información que se ofrece y los objetivos que se siguen, lo que explica que estos esfuerzos no estén teniendo el impacto esperado” (p. 89).

El autor indica también que la expectativa radica en que la difusión de resultados propiciará el aumento del aprendizaje de los estudiantes; en el mismo sentido apunta

Martinic (2010) cuando apunta que “[...] a puesta en marcha de los sistemas de evaluación tiene, como uno de sus supuestos, el hecho de que una vez conocidos los resultados, éstos serían ampliamente empleados por los usuarios en distintos ámbitos decisionales” (p. 37). Aunque la difusión de los resultados es una parte imprescindible de cualquier tipo de evaluación, no se reduce a la tarea de garantizar el acceso a la información, “[...] sino que deben ofrecerse las herramientas adecuadas para su interpretación, así como evitar usos inapropiados” (INEE, 2006, p. 1).

La evaluación a gran escala que actualmente opera en el sistema educativo mexicano, PLANEA, ha buscado, a través de su diseño e implementación, superar las limitaciones de las evaluaciones que la precedieron y mantener sus fortalezas (INEE, 20015). En los documentos que brindan el marco general de PLANEA, el INEE reconoce que la tarea de evaluar no se reduce a la mera producción de información por relevante que ésta sea, sino que se requiere conformar y consolidar procesos de difusión que garanticen el acceso oportuno de todos los actores educativos a la información, así como orientaciones que guíen su interpretación y análisis. De igual manera, se advierte que:

*Actualmente, uno de los grandes retos de las evaluaciones consiste en diseñarlas en función de modelos lógicos donde se prevean los usos que los distintos actores pueden dar a los resultados para producir los efectos deseados en el sistema educativo, al mismo tiempo que anticipar y desalentar usos inapropiados o no pertinentes.* (INEE, 2015, p. 2)

Ciertamente, los elementos involucrados en la difusión y uso de resultados de evaluaciones a gran escala, como muestra la revisión de literatura expuesta en este artículo, se constituyen como procesos cuya interacción es compleja, lo que hace indispensable la existencia de mecanismos que garanticen su congruencia, en aras de garantizar un adecuado aprovechamiento de la información que proporciona este tipo de evaluaciones.

En este sentido, también se torna imprescindible que las evaluaciones a gran escala sean objeto de estudio de investigaciones sistemáticas y rigurosas que examinen tanto su calidad técnica, implementación y procesamiento de datos, como las prácticas en torno a su análisis, uso y las consecuencias que derivan de ellos.

La propuesta que se hace en este artículo aporta los ejes que fueron de utilidad para orientar el acopio de información, así como la sistematización y análisis de ésta, en el marco de un estudio que tomó a la prueba ENLACE como su referente. Los ejes se refieren a procesos que abarcan un ciclo amplio una vez que los resultados están disponibles para hacerse públicos: difusión, análisis e interpretación, uso de los resultados propiamente, efecto del uso de éstos, y factores asociados al aprovechamiento de la información. Algunos se pudieron describir mejor que otros; pero en conjunto, sugieren vetas de indagación potencialmente relevantes para determinar si los esfuerzos y recursos que hoy en día se invierten en la evaluación de los resultados de aprendizaje están permitiendo alcanzar las expectativas que los alientan.

## **Financiación**

Proyecto financiado por el Consejo Nacional de Ciencia y Tecnología (CONACYT) y la Secretaría de Educación Pública (SEP).

## Referencias

- Aristimuño, A. y Kaztman, R. (2005). La evaluación de aprendizajes y su impacto en las políticas y las escuelas públicas de Uruguay. S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina*, (pp. 173-243). Santiago de Chile: PREAL.
- Backhoff, E. y Contreras, S. (2014). "Corrupción de la medida" e inflación de los resultados de ENLACE. *Revista Mexicana de Investigación Educativa*, 19(63), 1267-1283.
- Cervera, L., Martín, G. y Sánchez, C. (2008). Estudio georreferencial de la Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE) en el municipio de Juárez, Chihuahua: Análisis espacial. *Revista Electrónica de Investigación Educativa*, 10(1), 1-23.
- Council of the Great City Schools. (2012). *Charting success: Data Use and Student Achievement in Urban Schools*. Recuperado de [http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Charting\\_Success.pdf](http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Charting_Success.pdf)
- Cueto, S. (Ed.). (2005). *Uso e impacto de la información educativa en América Latina*. Santiago de Chile: PREAL.
- Dussel, I., Legarralde, M., Marcalain, G., Montes, N. y Pinkaz, D. (2005). Los usos de la información empírica en el sistema educativo. Un estudio de caso en tres jurisdicciones de la Argentina. En S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina*, (pp. 345-394). Santiago de Chile: PREAL.
- Fernández, T. y Midaglia, C. (2005). El uso de los informes generados por los sistemas de evaluación de aprendizaje en la educación primaria. Los casos de México y Uruguay. En S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina* (pp. 11-171). Santiago de Chile: PREAL.
- Figuroa, A. (2003). La perspectiva de maestros, directores y supervisores escolares sobre la evaluación externa. En M. Zorrilla (Coord.), *La evaluación de la educación básica en México 1990-2000. Una mirada a contraluz* (pp. 351-406). Ciudad de México: Universidad Autónoma de Aguascalientes.
- Griffin, P. (2009). Teacher's Use of Assessment Data. En C. Wyatt-Smith y J. J. Cumming (Eds.), *Educational Assessment in the 21<sup>st</sup> Century* (pp. 183-208). Amsterdam. Springer. [https://doi.org/10.1007/978-1-4020-9964-9\\_10](https://doi.org/10.1007/978-1-4020-9964-9_10)
- INEE. (2006). La difusión de los resultados de las evaluaciones. *Colección Folletos: Los temas de la evaluación*. Ciudad de México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/C/610/P1C610.pdf>
- INEE. (2015a). *Planea: una generación de pruebas. ¿Qué es PLANEA?*. Ciudad de México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P2/A/303/P2A303.pdf>
- Institute of Education Sciences. (2009). *Using Student Achievement Data to Support Instructional Decision Making*. Recuperado de [http://ies.ed.gov/ncee/wwc/pdf/practice\\_guides/dddm\\_pg\\_092909.pdf](http://ies.ed.gov/ncee/wwc/pdf/practice_guides/dddm_pg_092909.pdf)
- Love, N. (2004). Taking Data to New Depths. *National Staff Development Council*, 25(4), 22-26.
- Martínez Rizo, F. (2001). Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate. *Revista de la Educación Superior*, 30(120), 71-85.
- Martínez Rizo, F. (Coord.). (2015). *Las pruebas ENLACE y EXCALE. Un estudio de validación*. Ciudad de México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/C/148/P1C148.pdf>

- Martinic, S. (2010). La evaluación y las reformas educativas en América Latina. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 3(3), 30-43.
- Martínez Rizo, F. y Santos, A. (2009). Evaluación e investigación educativa. En A. Alba y R. Glazman (Coords.), *¿Qué dice la investigación educativa?* (pp. 265-304). Ciudad de México: COMIE.
- Matters, G. (2009). A Problematic Leap in the Use of Test Data: From Performance to Inference. En C. Wyatt-Smith, C. y J. J. Cumming (Eds.), *Educational Assessment in the 21<sup>st</sup> Century* (pp. 209-225). Amsterdam: Springer.
- Means, B., Padilla, C., DeBarger, A. y Bakia, M. (2009). *Implementing Data-Informed Decision Making in Schools-Teacher Access, Supports and Use*. Recuperado de <http://files.eric.ed.gov/fulltext/ED504191.pdf>
- Montoya, S., Cruz, J. y Vera, A. (2005). Evaluación de la calidad educativa: de los sistemas centrales al aula. Estudio del impacto de las políticas de devolución de información personalizada a las escuelas en tres provincias de Argentina. En S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina* (pp. 295- 343). Santiago de Chile: PREAL.
- Moreno, T. (2011). La cultura de la evaluación y la mejora de la escuela. *Perfiles Educativos*, XXXIII(31), 116-130.
- Protheroe, N. (2009). improving teaching and learning with data-based decisions: Asking the right questions and acting on the answers. *ERS Spectrum*, 19(3), 4-9.
- Ravela, P., Wolfe, R., Valverde, G. y Esquivel, J. (2000). *Los próximos pasos: ¿Hacia dónde y cómo avanzar en la evaluación de aprendizajes en América Latina?*. Santiago de Chile: PREAL-GRADE-GTEE.
- Ravela, P. (2001). *¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?* Santiago de Chile: PREAL.
- Ravela, P. (2010). *¿Qué pueden aportar las evaluaciones estandarizadas a la evaluación en el aula?* Documentos de Trabajo PREAL 47. Santiago de Chile: PREAL.
- Ravela, P. (2012). La experiencia del INEE en la difusión de resultados para docentes. En INEE, *INEE. Una década de Evaluación 2002-2012*. Ciudad de México: INEE.
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez, F., Aylwin, M. y Wolff, L. (2007). *Las evaluaciones educativas que América Latina necesita. Documentos de trabajo PREAL 40*. Santiago de Chile: PREAL.
- Ruiz, G. y Prado, P. (2017, noviembre). El uso en el nivel de educación primaria de los resultados de la evaluación ENLACE. Un estudio intensivo en Aguascalientes. Comunicación presentada en el *XIV Congreso Nacional de Investigación Educativa*, Consejo Mexicano de Investigación Educativa, San Luis Potosí. Recuperado de <http://comie.org.mx/congreso/memoriaelectronica/v14/doc/1037.pdf>
- Santiago, P., McGregor, I., Nusche, D., Ravela, P. y Toledo, D. (2014). *Revisiones de la OCDE sobre la Evaluación en Educación. México 2012*. Ciudad de México: INEE.
- SEP. (2013). *Evaluación Nacional del Logro Académico en Centros Escolares. Educación Básica: manual técnico 2013*. Ciudad de México: SEP.
- Young, V. y Kim, D.H. (2010). Using assessments for instructional improvement: A literature review. *Education Policy Analysis Archives*, 18(19), 1-39. <https://doi.org/10.14507/epaa.v18n19.2010>

## **Breve CV de las autoras**

### **Guadalupe Ruiz Cuéllar**

Doctora en Educación egresada del programa interinstitucional de la Universidad Autónoma de Aguascalientes. Es Investigadora Nivel 1 del Sistema Nacional de Investigadores. Es docente de pregrado y posgrado en las áreas de Metodología de la Investigación, Gestión de la Educación y Evaluación Educativa. Su trabajo de investigación se desarrolla fundamentalmente en esta última área. Actualmente es Directora General de Investigación y Posgrado de la UAA. Email: [gruiz@correo.uaa.mx](mailto:gruiz@correo.uaa.mx)

### **Paloma Prado Robledo**

Licenciada en Letras Españolas por la Universidad de Guanajuato y Maestra en Investigación Educativa por la Universidad Autónoma de Aguascalientes. Se ha dedicado al diseño y elaboración de materiales didácticos y evaluaciones de aprendizaje en contextos de educación para adultos. Fungió como formadora de profesores en diversos temas pedagógicos, fue coordinadora académica y docente en educación media superior. Actualmente, colabora en proyectos de investigación en la línea de práctica docentes y evaluación educativa. Email: [pradopaloma@hotmail.com](mailto:pradopaloma@hotmail.com)