

INVESTIGACIONES

Validación de un instrumento cuantitativo para medir la práctica reflexiva de docentes en formación*

Validating a quantitative instrument to measure reflective practice in student teachers

Álvaro Salinas,^a Eugenio Chandía,^b Daniela Rojas^{bc}

^aFacultad de Educación, Universidad Católica de Chile
Telf.: (56) 223545320. Correo electrónico: asalinase@uc.cl

^bPrograma de Doctorado en Educación, Facultad de Educación, Universidad Católica de Chile
Telf.: (56) 223545320. Correo electrónico: echandia@uc.cl

^cCorreo electrónico: dhrojas1@uc.cl

RESUMEN

La práctica reflexiva ha recibido un creciente interés de parte de la investigación y de la formación docente a nivel internacional. Sin embargo, hay poco acuerdo sobre qué es y cómo medirla. Este artículo se propone validar un instrumento cuantitativo en español, para medir la práctica reflexiva en estudiantes de pedagogía. Para ello se aplicó la traducción de la escala de reflexividad elaborada por Larrivee, a una muestra probabilística de 489 estudiantes de pedagogía en Chile. El análisis de datos incluyó análisis factorial exploratorio y confirmatorio y el cálculo de la confiabilidad de la escala. La escala resultante está compuesta por 20 ítems agrupados en tres dimensiones distintas (reflexión superficial, pedagógica y crítica), con buenos indicadores de confiabilidad. El trabajo contribuye a definir operacionalmente el concepto de práctica reflexiva y ofrece la posibilidad de medir el constructo. Al finalizar, se discuten limitaciones y proyecciones de este trabajo.

Palabras clave: formación de profesores, reflexión, validación de escala, práctica reflexiva.

ABSTRACT

Reflective practice has become a topic of increasing interest in teacher training and research circles worldwide. However, there is little consensus as to how it should be defined and measured. This paper aims to validate a quantitative instrument in Spanish to measure reflective practice in student teachers. To do so, a scale created by Larrivee to measure reflection was translated and administered to a probability sample of 489 student teachers in Chile. The data analysis portion of the study included both exploratory and confirmatory factor analyses, as well as a calculation of the scale's reliability. The resulting scale consists of 20 items grouped into three different dimensions (surface, pedagogical, and critical reflection), with good reliability indicators. This work offers an operational definition of reflective practice, as well as contributes with a tool to measure it. The article concludes by discussing the limitations of and prospects for the validated scale.

Key words: teacher education, reflection, scale validity, reflective practice.

* Este trabajo ha sido financiado por el Fondo Nacional de Desarrollo Científico y Tecnológico, Fondecyt, n° 1120834.

1. INTRODUCCIÓN

La práctica reflexiva ha recibido un creciente interés en el campo educativo, tanto de parte de la investigación como de programas de educación superior (Collin, Karsenti & Komis, 2012; Jaeger, 2013). El interés por la práctica reflexiva en educación se ha reflejado en una creciente producción académica, en esfuerzos por incorporar la práctica reflexiva en la formación de profesores y en su inclusión en estándares y perfiles profesionales.

En Hispanoamérica, varios sistemas educativos han incorporado la reflexividad como uno de los elementos clave del buen desempeño docente (Meckes, 2014). En Chile, Ecuador y Perú esto se ha hecho a través de estándares (Flotts & Abarzúa, 2011; Meckes, 2014; Ministerio de Educación de Perú, 2012). En Argentina se ha incorporado en los lineamientos para la formación docente inicial (Ministerio de Educación, 2007). En España, algunos programas de formación docente han incorporado la reflexividad como competencias de egreso (Escuela Universitaria de Magisterio Universidad Complutense de Madrid, s/f; Facultad de Educación de la Universidad Complutense de Madrid, s/f; Facultad de Educación de la Universidad de Castilla-La Mancha, 2015).

Durante las últimas décadas, la investigación sobre la reflexividad docente se ha ocupado de comprender qué es, cómo se desarrolla, cómo se mide y qué relación tiene con el mejoramiento de las prácticas docentes y el aprendizaje de los estudiantes (Beauchamp, 2014; Jaeger, 2013). Sin embargo, varios trabajos coinciden en señalar que, a pesar de los esfuerzos, aún hay poca claridad sobre estas cuestiones (Beauchamp, 2014; Jaeger, 2013; Lee, 2005).

Algunos autores señalan la necesidad de desarrollar investigación cuantitativa que permita complementar la investigación cualitativa dominante en este campo y enriquezca la comprensión de la práctica reflexiva (Collin et al., 2012; Jaeger, 2013; Standal & Moe, 2013).

En Hispanoamérica, muchos de los estudios más recientes tienden a analizar cualitativamente la reflexión docente, empleando para ello muestras constituidas por unos pocos programas de formación de profesores y/o un pequeño número de profesores o de estudiantes de pedagogía. En Chile, por ejemplo, el trabajo de González-Weil et al. (2014) estudió la contribución del trabajo colaborativo en la reflexión para la transformación de la práctica profesional en docentes de ciencias del sistema escolar y de educación superior. El estudio involucró una investigación acción realizada por 23 profesores. Para recoger información, el estudio incluyó cuestionarios, bitácoras, apuntes de observación y filmación de sesiones de trabajo. Fourés (2011), en Argentina, pidió a 8 profesores que seguían un curso de educación física en la universidad trabajar con las planificaciones de sus cursos y con “diarios de ruta” que contuvieran reflexiones sobre su propio proceso de aprendizaje. En Perú, Alberca y Frisancho (2011) estudiaron cuánto saben, cómo entienden y qué importancia asignan a la reflexión docente un grupo de 18 profesores de un establecimiento educacional de Ayacucho. Para ello usaron una encuesta con preguntas abiertas que luego fueron procesadas usando análisis de contenido. Castellanos y Yaya (2013), en Colombia, analizaron textos elaborados por 13 estudiantes que participaron en un curso de postgrado para profesores en ejercicio. Para el análisis, el equipo de investigadores empleó categorías generadas inductiva y deductivamente. En España, Mena y García (2013) analizaron informes de práctica elaborados por 104 estudiantes de pedagogía de una universidad. Los informes contenían la descripción y análisis que los estudiantes hicieron sobre incidentes críticos que ellos experimentaron en el contexto de su práctica profesional. El material fue analizado empleando técnicas de análisis de discurso.

Este trabajo busca aportar a la discusión sobre cómo medir la práctica reflexiva. Se propone específicamente validar un instrumento cuantitativo en castellano que permita medir la práctica reflexiva en estudiantes de pedagogía. Un instrumento como este puede complementar la investigación cualitativa realizada en este campo al incluir muestras mayores y con ello ofrecer una visión más general de las prácticas reflexivas en estudiantes de pedagogía, facilitar la comparación de programas de formación y analizar la relación entre la práctica reflexiva y otras variables.

1.1. EL CONCEPTO DE PRÁCTICA REFLEXIVA

Para efectos de este trabajo, se entiende la práctica reflexiva como el examen crítico y cotidiano del trabajo docente a partir de marcos interpretativos (“appreciative systems”, en términos de Schön, 1993), los que pueden poner su foco en distintos aspectos de la práctica examinada: situaciones específicas de la docencia en la sala de clases, el marco institucional donde trabaja el profesor o el sistema escolar y social que encuadra y orienta la tarea del docente (Russell, 2005; Schön, 1983). El examen se produce en un proceso de enmarcamiento y reenmarcamiento (“framing” y “reframing”, en Schön), los que permiten dar orden y comprender situaciones complejas (“puzzling situations”, en Schön) (Loughran, 2002; Schön, 1983).

Algunos autores, inspirados en buena medida por van Manen (1977), han considerado que la reflexión docente se organiza en niveles de menor a mayor complejidad (Astudillo, Rivarosa & Ortiz, 2014; Mena & García, 2013; Wunder, 2003). Poniendo distancia de estos planteamientos, un conjunto de trabajos plantea que lo que usualmente se ha tratado como niveles son más bien tipos de reflexión (distintos marcos interpretativos) que son útiles para analizar en toda su complejidad el trabajo del profesor y los contextos en los cuales este se produce (Collin et al., 2012; Leijen et al., 2012).

1.2. LA MEDICIÓN CUANTITATIVA DE LAS PRÁCTICAS REFLEXIVAS

Entre los estudios que analizan cuantitativamente la práctica reflexiva se pueden distinguir aquellos que emplean instrumentos que analizan el desempeño de estudiantes de pedagogía o de profesores en ejercicio y aquellos que recogen el reporte que los propios sujetos hacen sobre sus prácticas reflexivas (Kalk et al., 2014).

El trabajo de Concha, Hernández, del Río, Romo y Andrade (2013) es un ejemplo del primer tipo de estudio. Los autores analizan en Chile la capacidad de los estudiantes de pedagogía en educación básica para reflexionar pedagógicamente. Para ello emplean una muestra de 82 sujetos de tres instituciones formadoras, a los que se les mostró un video de clases, luego de lo cual debían escribir un análisis. Los textos elaborados eran analizados según una rúbrica con cinco niveles (uno alto, tres intermedios y uno bajo). Para asignar a los estudiantes a cada nivel, la rúbrica empleaba cuatro criterios: la inferencia de las razones que subyacen a las decisiones del profesor observado en el video, la articulación de distintos dominios de conocimiento relevantes a la enseñanza y el aprendizaje, el dominio de conceptos teóricos empleados y un análisis fundamentado.

El trabajo de Kalk et al. (2014) es un ejemplo de estudio del segundo tipo. En él los autores analizan la validez y confiabilidad de dos escalas que recogen el reporte que los propios sujetos hacen de sus prácticas reflexivas en procesos de aprendizaje. Para ello

aplican los instrumentos a una muestra de 216 estudiantes de pedagogía en Estonia. Una de las escalas empleadas fue desarrollada por Kember et al. (2000); la otra fue desarrollada por Sobral (2001). Ambas escalas fueron inicialmente desarrolladas para ser empleadas con estudiantes de ciencias de la salud. La escala de Kember incluye 16 ítems que miden 4 sub-dimensiones: acciones habituales, comprensión, reflexión y reflexión crítica. Cada sub-dimensión es medida con 4 ítems, que se responden usando una escala Likert de 5 puntos. La escala de Sobral es unidimensional e incluye 14 ítems, los que son respondidos en una escala de 7 puntos que miden la frecuencia en la que quienes responden realizan ciertas prácticas de aprendizaje reflexivo. Los resultados del estudio de Kalk et al. (2014) muestran una confiabilidad moderada de ambas escalas, probablemente por la pequeña cantidad de ítems de cada escala.

Otro trabajo que desarrolló una escala para medir prácticas reflexivas es el de Larrivee (2008). Esta escala fue originalmente desarrollada para el campo de la educación, a diferencia de las escalas empleadas por Kalk et al. (2014). Su propósito fue el de ofrecer una herramienta de medición que pueda ser usada para identificar el nivel de reflexión de un estudiante de pedagogía o de un profesor en ejercicio.

La escala de Larrivee distingue tres niveles de reflexión docente. Un primer nivel, de reflexión superficial, se define como el examen de aspectos “tácticos” de la enseñanza. En este nivel se examinan cuestiones tales como las mejores formas para alcanzar objetivos y estándares predefinidos. Las creencias y posturas del profesor se sustentan en la experiencia, pero no en resultados de investigación ni teoría. El profesor identifica diferencias entre sus estudiantes y es capaz de acomodarse a ellas. Un segundo nivel, de reflexión pedagógica, implica un examen de la práctica docente a la luz de la teoría educacional, de resultados de investigación y de la experiencia. El profesor está constantemente pensando en cómo sus prácticas de enseñanza afectan el aprendizaje de los alumnos y en cómo expandir las experiencias de aprendizaje. Un tercer nivel, de reflexión crítica, involucra un análisis de las condiciones de posibilidad y las consecuencias éticas, sociales y políticas de sus prácticas (Larrivee, 2008).

Además de estos tres niveles de práctica reflexiva, Larrivee agrega un nivel inicial, llamado pre-reflexivo, según el cual aquellas cosas que ocurren durante la enseñanza son consideradas por el profesor como eventos aislados, sin una atención atenta de su conexión con otros eventos o circunstancias. Estos eventos se consideran fuera de su control. En este nivel la reacción del profesor es reactiva y no se funda en evidencia o experiencia (Larrivee, 2008).

La escala de Larrivee tiene tres atributos importantes. El primero es el que incluye un número de ítems que permite un buen examen de sus atributos de confiabilidad y validez estadística. El segundo es que los ítems fueron elaborados y clasificados por jueces expertos. El tercero y más importante es que ha sido desarrollada específicamente para las prácticas de reflexión docente. Considerando estos atributos, la escala que ha sido sometida a validación es la de Larrivee, traducida al castellano.

2. MÉTODO

Para construir los ítems de la escala, Larrivee elaboró un instrumento que contenía una descripción general de niveles de reflexividad, así como ítems específicos elaborados a partir de la literatura revisada. Este instrumento fue enviado a investigadores que habían

trabajado o escrito recientemente en el tema, y que contribuyeron como jueces expertos clasificando los ítems en cada uno de los niveles de reflexividad. De las respuestas obtenidas se retuvieron 53 ítems que fueron clasificados más claramente en cada nivel por la mayor cantidad de los 40 jueces expertos que respondieron a la encuesta: 14 de los ítems corresponden a los niveles de pre-reflexión, reflexión pedagógica y reflexión crítica, y 11 corresponden al nivel de reflexión superficial.

Considerando que el foco del trabajo que aquí se presenta son las prácticas reflexivas, se eliminó del análisis el nivel de pre-reflexión, elaborado por Larrivee para dar cuenta de sujetos que están en un nivel previo y que no alcanzan a desarrollar prácticas reflexivas. De este modo, lo que sigue incluye solo 39 ítems que corresponden a los niveles de reflexión superficial, pedagógica y crítica.

El instrumento definitivo elaborado por Larrivee tiene dos versiones. Una tiene como propósito que los estudiantes de pedagogía puedan auto-evaluar la frecuencia con la que despliegan distintas prácticas reflexivas. La otra tiene como propósito que un formador de profesores, supervisor o mentor describa la frecuencia de esas prácticas en profesores en formación. La versión que ha sido considerada en este trabajo es la de auto-evaluación.

La escala original está en idioma inglés. Para su uso en un contexto de habla castellana, el instrumento fue traducido por un traductor experto. Algunos ítems fueron modificados para ajustarse al contexto social y cultural de aplicación. Posteriormente, otro traductor hizo una traducción en sentido contrario (castellano a inglés), de modo de revisar el ajuste de la primera traducción.

La escala original de Larrivee admite tres alternativas de respuesta: “frecuentemente”, “a veces” y “raramente”. La escala validada aquí incorpora otras dos. Una para estudiantes de pedagogía que hayan tenido la oportunidad de desarrollar sus prácticas reflexivas durante su formación, pero no las aprovecharon; y otra para quienes no hayan tenido la oportunidad de desarrollar este tipo de prácticas durante su formación. Las alternativas empleadas reciben una puntuación que va de 5 a 1 punto.

Antes de la aplicación masiva de este instrumento, se solicitó a un pequeño grupo de 13 estudiantes de pedagogía que respondieran la escala, de modo de estimar la duración de la aplicación, la claridad de cada ítem y la facilidad de respuesta. A partir de sus comentarios, algunos ítems fueron modificados en su redacción para mejorar su comprensión.

2.1. MUESTRA

La muestra empleada es probabilística. El instrumento fue aplicado a 650 estudiantes de pedagogía que estaban cursando su último año de estudios en alguno de los 134 programas de formación de profesores primarios y secundarios seleccionados.

Estos programas de formación eran impartidos por 37 instituciones formadoras de profesores, ubicadas en cinco de las regiones con mayor matrícula en programas de formación de profesores de Chile.

La selección de los participantes se realizó en dos etapas. En la primera, se seleccionaron al azar los programas de formación de profesores, para lo cual se empleó la base de datos de programas de formación de profesores del año 2013 del Consejo Nacional de Educación de Chile (Consejo Nacional de Educación, 2014). Una vez seleccionados los programas, se contactó por correo electrónico o teléfono a los directivos de los programas para solicitar su apoyo en el estudio. En los casos de rechazo, el programa fue reemplazado.

En la segunda etapa se seleccionó de forma aleatoria simple a cinco estudiantes de cada programa de formación que estuvieran presentes en clases el día de la aplicación. Esta aplicación era presencial. Una vez que cada participante firmaba el consentimiento informado que se le entregaba por escrito a cada uno, encuestador les proporcionaba un cuestionario en papel, que era contestado en alrededor de 30 minutos.

La muestra obtenida se compone de 58% de mujeres y 42% hombres, con una edad media de 24,6 años (D.E.=4.5). El universo de estudiantes de pedagogía matriculados en programas de formación de profesores en las cinco regiones seleccionadas se compone de 55% de mujeres y 45% hombres. Según estos datos, la muestra tiene un buen ajuste respecto al universo de referencia, aunque las mujeres están ligeramente sobrerrepresentadas. Esto puede responder a que el universo de referencia abarca a todos los estudiantes matriculados en los programas de formación de profesores, y no solo a los estudiantes de último año, como fue el caso de la muestra obtenida.

Para el análisis de datos se eliminaron los cuestionarios que no estuvieran completos, de manera que finalmente se trabajó con una base de 489 casos. Este número de casos es adecuado para los análisis estadísticos previstos (MacCallum et al., 1999).

2.2. ANÁLISIS DE DATOS

El análisis de datos se realizó en cuatro etapas. La primera incluye una correlación ítem test, que busca determinar los ítems que se relacionan o no con la escala. La segunda incluye un análisis factorial exploratorio, que busca identificar los factores latentes que permiten organizar los ítems que componen la escala. En la tercera etapa se realiza análisis factorial confirmatorio, el cual se centra en reducir la cantidad de indicadores, conservando la mayor varianza. Por último, en la cuarta etapa se determina la consistencia interna de las sub-escalas identificadas.

En la primera etapa se utiliza el coeficiente “r” de Pearson. Los ítems que comparten menos varianza con el grupo son analizados cuidadosamente y eliminados si la correlación es menor a 0,4. En la segunda etapa se emplea un análisis factorial exploratorio (AFE) para determinar si los indicadores propuestos en el instrumento comparten alguna estructura latente. Luego de levantar los factores con AFE, en la tercera etapa se realiza un análisis factorial confirmatorio (AFC), que tiene por objetivo la búsqueda de combinaciones de variables observadas (ítems de la escala) en un número inferior de variables latentes, como lo son los constructos de reflexión que organizan los ítems. Por último, en la cuarta etapa se realiza un análisis de confiabilidad, determinando los coeficientes de consistencia interna para cada una de las variables no observadas. Los coeficientes de consistencia que se van a determinar son las cotas determinadas por Guttman (1945), entre las cuales se encuentra el coeficiente Alfa de Cronbach.

Los índices estadísticos empleados para determinar el ajuste de los modelos AFE y AFC son el absoluto, el índice relativo Tucker-Lewis (TLI) y el índice comparativo de ajuste (CFI). Por último, se estimará el Error Cuadrático Medio de Aproximación (RMSEA). Para el ajuste de los modelos, se han considerado como aceptables valores TLI y CFI mayores a 0,9. Valores mayores 0,95 se han considerado como buenos (Hu & Bentler, 1999). Por último, para RMSEA se han considerado valores menores a 0,05 como buen ajuste, y entre 0,05 y 0,08 como un ajuste regular (Schermelleh-engel, Moosbrugger & Muller, 2003).

3. RESULTADOS

3.1. CORRELACIÓN ÍTEM TEST

Para analizar la consistencia interna del instrumento se calculó la correlación entre cada ítem y el puntaje total de la escala. Los resultados que aparecen en la Tabla 1 muestran 7 ítems con correlaciones bajo 0,4, los que fueron eliminados. Los ítems eliminados fueron los V6, V10, V13, V18, V19, V32, V38).

Tabla 1. Correlaciones ítem-test

N° Ítem	Ítem	Nivel reflex.	Correlación ítem-test
V1	ajusta las metodologías y las estrategias según el rendimiento de los estudiantes	pedag.	,591**
V2	realiza pocos cambios en la metodología de enseñanza o evaluación según los distintos estilos de aprendizaje de los alumnos	super.	,439
V3	considera en el proceso de aprendizaje las experiencias y conocimientos previos que el estudiante trae consigo	pedag.	,673**
V4	pone en tela de juicio creencias sobre educación comúnmente aceptadas por los profesores	critic.	,431
V5	contribuye a solucionar problemas de equidad y justicia social que afectan a los estudiantes	critic.	,532*
V6	considera que basta con su experiencia para fundamentar sus creencias	super.	,323
V7	busca patrones, relaciones y conexiones para profundizar la comprensión de la situación de aprendizaje	pedag.	,672**
V8	modifica las estrategias de enseñanza cuestionando los supuestos implícitos sobre la enseñanza y el aprendizaje	super.	,585*
V9	tiene en cuenta que la enseñanza puede contribuir o bien impedir la realización de una sociedad más justa y humana	critic.	,574*
V10	tiene dificultades para conectar métodos de enseñanza con teorías implícitas	super.	,244
V11	se esfuerza por lograr el aprendizaje para todos los estudiantes	pedag.	,704***
V12	se esfuerza por experimentar y asumir riesgos para mejorar la efectividad de las prácticas de enseñanza	pedag.	,700**
V13	limita el análisis de las prácticas de enseñanza solo a cuestiones técnicas	super.	,269

V14	busca maneras de conectar los nuevos conceptos con el conocimiento previo de los estudiantes	pedag.	,725***
V15	analiza el impacto que tienen las distintas metodologías de trabajo en clases (tales como grupos de aprendizaje cooperativo, trabajo entre pares u otras) en el aprendizaje de los alumnos,	pedag.	,626**
V16	deja de lado algunos juicios para considerar todas las opciones posibles	critic.	,526*
V17	realiza bastantes cambios en la metodología de enseñanza o evaluación para acoger las diferencias individuales de los estudiantes	super.	,613**
V18	se centra en los aspectos específicos de una clase (metodologías, recursos, actividades, etc.) sin mirar el conjunto de la unidad u objetivos de aprendizajes más a largo plazo	super.	,306
V19	ajusta las prácticas de enseñanza solo a las situaciones actuales, sin desarrollar un plan a largo plazo	super.	,338
V20	cuando hay una incongruencia entre sus creencias y sus acciones, toma medidas para rectificar	critic.	,480*
V21	considera atentamente las conexiones entre las acciones del profesor y el aprendizaje de los estudiantes	pedag.	,649**
V22	se hace autocríticas constructivas sobre su propia enseñanza	pedag.	,591*
V23	reconoce la complejidad de la dinámica del aula	pedag.	,583*
V24	considera el punto de vista de los estudiantes en la toma de decisiones	pedag.	,668**
V25	es un investigador activo en la sala de clases, cuestionando conclusiones actuales y generando nuevas hipótesis	critic.	,597*
V26	identifica los supuestos y premisas que subyacen a su forma de ver la educación	critic.	,582*
V27	considera las implicaciones éticas de las prácticas en el aula	critic.	,583*
V28	favorece que sus estudiantes se hagan responsables de las implicancias sociales de sus acciones	critic.	,586*
V29	considera que la investigación puede aportar a la enseñanza	pedag.	,569*
V30	hace ajustes a su docencia basados en su experiencia pasada (exitosa o no exitosa)	super.	,610**
V31	reconoce las consecuencias sociales y políticas de su propia enseñanza	critic.	,539*

V32	no logra reconocer patrones comunes en las diferentes acciones de sus estudiantes	super.	,249
V33	cuestiona los supuestos sobre los estudiantes y las expectativas (positivas o negativas) que hay sobre ellos	critic.	,510*
V34	tiene compromiso con el aprendizaje permanente y la mejora de la práctica	pedag.	,668**
V35	implementa soluciones a problemas que ponen su foco en resultados de corto plazo	super.	,535**
V36	considera la práctica dentro de un contexto sociológico, cultural, histórico y político más amplio	critic.	,433*
V37	analiza su propia forma de pensar	critic.	,497*
V38	cuestiona la situación actual de normas y prácticas educativas respecto al poder y el control en la sala de clases	critic.	,357*
V39	conoce formas alternativas de representar ideas y conceptos para enseñar a los estudiantes	pedag.	,519*

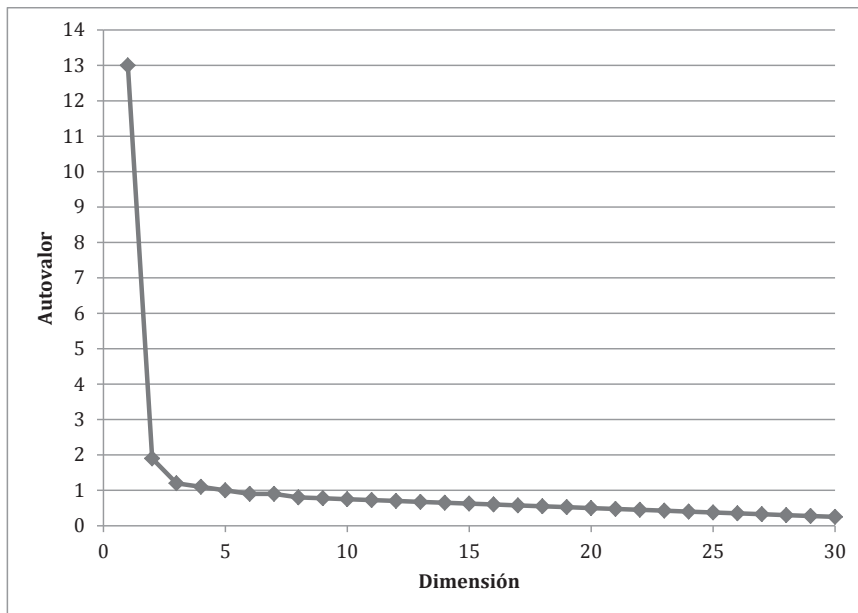
3.2. ANÁLISIS FACTORIAL EXPLORATORIO

La muestra y sus características permiten realizar análisis factorial exploratorio y confirmatorio. El valor de Kaiser-Meyer-Olkin (KMO) de adecuación muestral fue de 0,96, lo que significa que hay factores comunes que pueden estar explicando esta alta correlación.

Por otra parte, el test Bartlett's de esfericidad es significativo al 99%, con $\chi^2(496, n=489) = 7687,502$, $p < ,05$, lo que significa que la matriz de correlaciones de los datos observados, es diferente a la matriz identidad. Esto confirma que los ítems de la escala de reflexión se pueden reducir a variables latentes.

El Gráfico 1 muestra la sedimentación de factores. Como se aprecia, hay entre 4 y 6 factores con autovalores cercanos a 1, lo que indica que es posible encontrar ese mismo número de factores o dimensiones latentes en la escala.

Gráfico 1. Sedimentación de dimensiones (escala completa)



Considerando los datos del Gráfico 1, se examinará la extracción de 4, 5 y 6 factores. Para realizar el AFE, se consideran los 32 ítems retenidos luego de la correlación ítem-test, con la forma de extracción “ML” y rotación “VARIMAX”. En la Tabla 2 se observa que el modelo de 6 factores es el que tiene mejores ajustes, tanto absolutos, con $\chi^2(319, n=489) = 540,35, p < ,001$, como relativos (TLI=0,952; RMSEA=0,039). Sin embargo, el último solo explica un 3% de la varianza. Por otra parte, el valor propio asociado es mayor a 1 por solo 0,11 unidades. Esto aconseja descartar el modelo con 6 factores. De los modelos con 4 y 5 factores, el que presenta mejor ajuste es el segundo, con $\chi^2(346, n=489) = 637,41, p < ,001$ y TLI=0,941; RMSEA=0,043.

Tabla 2. Resumen de las estimaciones realizadas para las extracciones de 4, 5 y 6 factores

F.	χ^2 (gl)	RMSEA	TLI	VARIANZA EXPLICADA POR FACTOR						VALOR PROPIO POR FACTOR					
				F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6
4	755,64 (374) **	0,047	0,929	0,15	0,13	0,11	0,07			4,89	4,29	3,43	2,26		
5	637,41 (346) **	0,043	0,941	0,13	0,09	0,09	0,09	0,08		4,05	3,04	2,91	2,72	2,71	
6	540,35 (319) **	0,039	0,952	0,13	0,09	0,09	0,09	0,06	0,03	4,10	3,03	2,87	2,80	1,99	1,11

Nota. *p<0,05, **p<0,01.

Al analizar las cargas de los ítems para el modelo de 5 factores, se obtienen 5 ítems que cargan de forma cruzada en factores que más claramente no respetan la formulación teórica (Tabla 3). Estos ítems son el V1, que carga en el factor ML2 (con ítems con mayor carga en reflexión superficial), de reflexión pedagógica; el ítem V16 que carga en el factor ML5 (con mayoría de ítems de reflexión pedagógica), de reflexión crítica; el ítem V25 que carga en el factor ML3, que aglutina a indicadores de las tres sub-escalas; el ítem V30 que carga en ML1, que incluye ítems de las tres dimensiones, de reflexión superficial; y el ítem V39, que al igual que ítem V25 cae el factor ML3, que incluye a las tres sub-escalas. Al eliminar los ítems anteriores, la escala se reduce a 27.

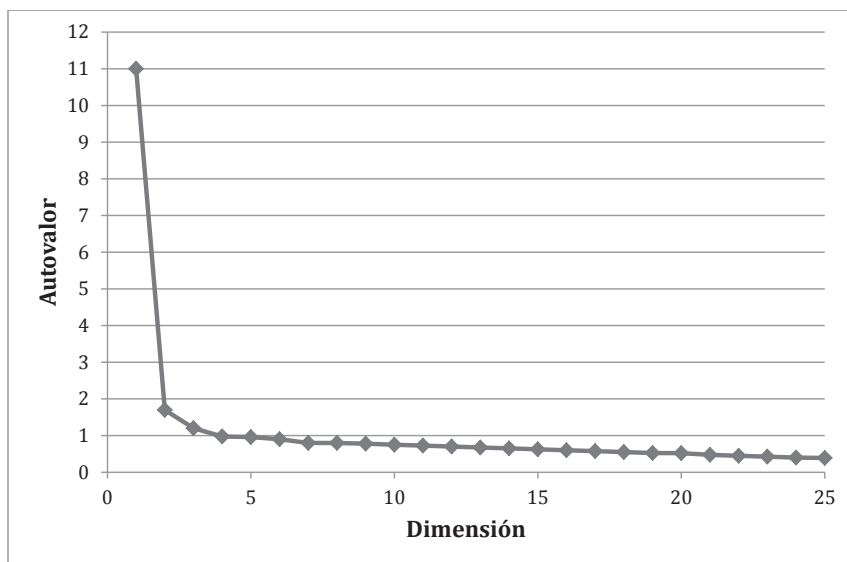
Tabla 3. Cargas factoriales, modelo con 5 factores

Ítem	Escala original	ML1	ML2	ML3	ML4	ML5
V1	pedag.	0,21	0,75	0,16	0,18	0,13
V2	super.	0,09	0,54	0,12	0,06	0,19
V3	pedag.	0,26	0,33	0,14	0,28	0,59
V4	critic.	0,04	0,10	0,13	0,54	0,16
V5	critic.	0,08	0,24	0,37	0,35	0,24
V7	pedag.	0,22	0,29	0,35	0,35	0,41
V8	super.	0,08	0,36	0,40	0,25	0,32
V9	critic.	0,36	0,10	0,09	0,37	0,39
V11	pedag.	0,53	0,29	0,27	0,05	0,53
V12	pedag.	0,29	0,32	0,48	0,14	0,43
V14	pedag.	0,41	0,29	0,22	0,16	0,62
V15	pedag.	0,19	0,30	0,40	0,18	0,42
V16	critic.	0,25	0,14	0,15	0,28	0,45
V17	super.	0,14	0,59	0,39	0,13	0,25
V20	critic.	0,20	0,26	0,08	0,39	0,12
V21	pedag.	0,43	0,29	0,16	0,33	0,32
V22	pedag.	0,56	0,25	0,09	0,20	0,30
V23	pedag.	0,58	0,16	0,12	0,28	0,20
V24	pedag.	0,43	0,19	0,40	0,21	0,34
V25	critic.	0,30	0,21	0,65	0,18	0,12
V26	critic.	0,26	0,18	0,43	0,41	0,09
V27	critic.	0,36	0,21	0,17	0,50	0,17
V28	critic.	0,33	0,23	0,28	0,35	0,17
V29	pedag.	0,58	0,10	0,33	0,23	0,11
V30	super.	0,51	0,26	0,19	0,20	0,28

V31	critic.	0,46	0,02	0,23	0,42	0,12
V33	critic.	0,24	0,19	0,14	0,36	0,20
V34	pedag.	0,62	0,19	0,29	0,24	0,23
V35	super.	0,23	0,36	0,36	0,17	0,14
V36	critic.	0,30	0,01	0,16	0,50	0,04
V37	critic.	0,41	0,03	0,10	0,44	0,15
V39	pedag.	0,29	0,13	0,35	0,28	0,17

El Gráfico 2 muestra la sedimentación para la escala reducida a 27 ítems. En él se observa la existencia de solo 3 factores con autovalores mayores a 1. Un cuarto factor está ligeramente bajo el 1.

Gráfico 2. Sedimentación de dimensiones (escala de 27 ítems)



Un nuevo cálculo de estimaciones para 3 y 4 factores, pero con la base con 27 ítems, se muestra en la Tabla 4. En ella se aprecia que el modelo de 4 factores es el que presenta mejores ajustes absolutos y relativos, con $\chi^2(249, n=489) = 409,38, p < ,001$ y TLI=0,961; RMSEA=0,038. Sin embargo, el cuarto factor solo explica un 6% de la varianza total.

Tabla 4. Resumen de las estimaciones realizadas para las extracciones de 3 y 4 factores para 27 ítems

Fac.	χ^2 (gl)	RMSEA	TLI	VARIANZA EXPLICADA POR FACTOR				VALOR PROPIO POR FACTOR			
				F1	F2	F3	F4	F1	F2	F3	F4
3	499,82 (273)**	0,042	0,95	0,17	0,15	0,13		4,67	4,03	3,54	
4	409,38 (249)**	0,038	0,961	0,16	0,15	0,10	0,06	4,41	3,99	2,61	1,68

Nota. *p<0,05, **p<0,01.

Para el modelo de 4 factores se hizo un nuevo cálculo de las cargas de los ítems. La Tabla 5 muestra que dos de los ítems presentan cargas en factores que no corresponden al constructo original. El ítem V5 carga en el factor del constructo reflexión superficial, pero pertenece teóricamente al de reflexión crítica. El ítem V9 carga en el factor del constructo reflexión pedagógica, pero al igual que el V5, teóricamente fue definido en el constructo de reflexión crítica. Al eliminar ambos, la escala queda compuesta por 25 ítems. Con ellos se procederá a realizar un análisis factorial confirmatorio.

Tabla 5. Cargas factoriales, modelo con 4 factores

Ítem	Escala original	ML1	ML2	ML3	ML4
V2	super.	0,19	0,04	0,46	0,09
V3	pedag.	0,53	0,17	0,47	0,31
V4	critic.	0,09	0,19	0,16	0,57
V5	critic.	0,06	0,25	0,5	0,3
V7	pedag.	0,29	0,3	0,52	0,3
V8	super.	0,17	0,15	0,58	0,26
V9	critic.	0,4	0,36	0,2	0,29
V11	pedag.	0,59	0,37	0,49	-0,01
V12	pedag.	0,32	0,3	0,65	0,07
V14	pedag.	0,62	0,28	0,47	0,14
V15	pedag.	0,29	0,2	0,57	0,17
V17	super.	0,17	0,17	0,69	0,11
V20	critic.	0,11	0,29	0,26	0,31
V21	pedag.	0,41	0,4	0,33	0,25
V22	pedag.	0,48	0,46	0,24	0,09

V23	pedag.	0,4	0,52	0,17	0,17
V24	pedag.	0,35	0,42	0,44	0,15
V26	critic.	0,08	0,44	0,37	0,27
V27	critic.	0,16	0,5	0,29	0,33
V28	critic.	0,15	0,41	0,38	0,24
V29	pedag.	0,25	0,59	0,25	0,07
V31	critic.	0,15	0,58	0,18	0,23
V33	critic.	0,22	0,29	0,23	0,32
V34	pedag.	0,33	0,63	0,35	0,03
V35	super.	0,08	0,31	0,52	0,06
V36	critic.	0,04	0,49	0,11	0,35
V37	critic.	0,25	0,49	0,06	0,3

3.3. ANÁLISIS FACTORIAL CONFIRMATORIO

El modelo de 3 factores estimado con AFE es el siguiente: el primer factor lo conforman 4 ítems originalmente asociados a reflexión superficial, más 4 originalmente asociados a la reflexión pedagógica (Tabla 6); el segundo lo conforman 8 ítems de reflexión pedagógica, y el tercero lo conforman 9 ítems originalmente asociados a reflexión crítica.

El modelo de 4 factores estimado con AFE es el siguiente: el primer factor lo conforman 4 ítems de reflexión superficial, más 4 de reflexión pedagógica; el segundo factor se compone de 6 ítems de reflexión crítica y 3 de reflexión pedagógica; el tercer factor se compone de 5 ítems de reflexión pedagógica, y el cuarto factor se compone de 3 ítems de reflexión crítica.

Tabla 6. Ítems en modelos de 3 y 4 factores

Modelo	Factor		Ítem
3	1	Reflexión Superficial	2, 8, 17 y 35
		Reflexión Pedagógica	3, 7, 12 y 15
	2	Reflexión Pedagógica	11, 14, 21, 22, 23, 24, 29 y 34
	3	Reflexión Crítica	4, 20, 26, 27, 28, 31, 33, 36 y 37
4	1	Reflexión Superficial	2, 8, 17 y 35
		Reflexión Pedagógica	7, 12, 15 y 24
	2	Reflexión Crítica	26, 27, 28, 31, 36 y 37
		Reflexión Pedagógica	23, 29 y 34
	3	Reflexión Pedagógica	3, 11, 14, 21 y 22
	4	Reflexión Crítica	4, 20 y 33

Al realizar las estimaciones con ambos modelos (Tabla 7), se aprecia que el modelo con tres factores es el que tiene mejores ajustes relativos (CFI=0,923; TLI=0,915; RMSEA=0,057) y absolutos, con $\chi^2(249, n=489) = 642,395, p < ,001$. Así, se considera entonces que la escala reflexión presenta 3 constructos latentes que se asocian a los teóricos definidos y conforman tres sub-escalas.

Tabla 7. Resumen de las estimaciones realizadas para las extracciones de 3 y 4 factores con 25 ítems

Modelo	χ^2 (gl)	CFI	TLI	RMSEA	RMSEA (90 CI)
3 factores	642,395 (249)**	0,923	0,915	0,057	0,051 – 0,062
4 factores	707,172 (269)**	0,920	0,911	0,058	0,053 – 0,063

Nota. *p<0,05 **p<0,01.

Sin embargo, en el factor 1 del modelo de 3 factores existen indicadores de dos constructos teóricos diferentes, de reflexión superficial y pedagógica (Tabla 6). Luego, para respetar la escala teórica, se han eliminado los ítems de reflexión pedagógica. Además se han eliminado aquellos ítems donde los factores latentes no cargan sobre 0,5, los que corresponden al ítem V2 en el factor de reflexión superficial y al ítem V4 del factor de reflexión crítica, lo que ha mejorado los ajustes de todos los índices como se observa en la Tabla 8.

Tabla 8. Resumen de las estimaciones realizadas para 3 factores

Modelo		CFI	TLI	RMSEA	RMSEA (90 CI)
3 factores	376,205 (149)**	0,941	0,932	0,056	0,049 – 0,063

Nota. *p<0,05, **p<0,01.

De esta forma, las cargas factoriales de cada uno de los ítems son los que presenta la Tabla 9.

Tabla 9. Cargas factoriales para el modelo de 3 factores

Factor	No Ítem	Ítem	Carga
F1: reflexión superficial	V8	Modifica las estrategias de enseñanza cuestionando los supuestos implícitos sobre la enseñanza y el aprendizaje	0,672***
	V17	Realiza bastantes cambios en la metodología de enseñanza o evaluación para acoger las diferencias individuales de los estudiantes	0,739***
	V35	Implementa soluciones a problemas que ponen su foco en resultados de corto plazo	0,592***
F2: reflexión pedagógica	V11	Se esfuerza por lograr el aprendizaje para todos los estudiantes	0,795***
	V14	Busca maneras de conectar los nuevos conceptos con el conocimiento previo de los estudiantes	0,766***
	V21	Considera atentamente las conexiones entre las acciones del profesor y el aprendizaje de los estudiantes	0,712***
	V22	Se hace autocríticas constructivas sobre su propia enseñanza	0,697***
	V23	Reconoce la complejidad de la dinámica del aula	0,673***
	V24	Considera el punto de vista de los estudiantes en la toma de decisiones	0,734***
	V29	Considera que la investigación puede aportar a la enseñanza	0,647***
	V34	Tiene compromiso con el aprendizaje permanente y la mejora de la práctica	0,756***
F3: reflexión crítica	V20	Cuando hay una incongruencia entre sus creencias y sus acciones, toma medidas para rectificar	0,515***
	V26	Identifica los supuestos y premisas que subyacen a su forma de ver la educación	0,629***
	V27	Considera las implicaciones éticas de las prácticas en el aula	0,702***
	V28	Favorece que sus estudiantes se hagan responsables de las implicancias sociales de sus acciones	0,627***
	V31	Reconoce las consecuencias sociales y políticas de su propia enseñanza	0,639***
	V33	Cuestiona los supuestos sobre los estudiantes y las expectativas (positivas o negativas) que hay sobre ellos	0,530***
	V36	Considera la práctica dentro de un contexto sociológico, cultural, histórico y político más amplio	0,540***
	V37	Analiza su propia forma de pensar	0,572***

Nota. *p<0,05, **p<0,01.

Las correlaciones entre los constructos varían entre $r(3)=0,348$, $p < ,001$ y $r(3)=0,426$, $p < ,001$. La Tabla 10 muestra correlaciones entre los factores.

Tabla 10. Correlaciones de los factores

	F2	F3
F1	0,426**	0,348**
F2		0,403**

Nota. * $p < 0,05$, ** $p < 0,01$.

3.4. CONSISTENCIA INTERNA

La escala completa tiene una consistencia interna que varía entre 0,89 como subestimación y 0,95 como sobre-estimación de la confiabilidad (Tabla 11). Para la escala completa el Alfa de Cronbach alcanza a 0,93. La sub-escala de reflexión superficial tiene una confiabilidad que varía entre 0,47 y 0,71. Esta sub-escala es la que presenta menor consistencia interna, lo que también se observa en la menor carga factorial a cada uno de los indicadores que la componen. La segunda sub-escala de reflexión pedagógica presenta una confiabilidad que varía entre 0,78 y 0,92 siendo la más alta entre las sub-escalas. Por último, la sub-escala de reflexión crítica tiene una confiabilidad que varía entre 0,71 y 0,84.

Tabla 11. Consistencia interna de cada factor y de la escala en su conjunto

Factor Reflexión	Ítem	VALORES LAMBDA DE GUTTMAN PARA CONSISTENCIA INTERNA					
F1: Superficial	V8, V17, V35	0,47	0,7	0,7	0,69	0,71	0,62
F2: Pedagógica	V11, V14, V21, V22, V23, V24, V29, V34	0,78	0,9	0,9	0,92	0,87	0,89
F3: Crítica	V20, V26, V27, V28, V31, V33, V36, V37	0,71	0,82	0,81	0,84	0,79	0,8
Escala Completa		0,89	0,93	0,93	0,95	0,91	0,94

4. DISCUSIÓN Y CONCLUSIONES

El propósito de este artículo ha sido validar una escala en castellano para medir prácticas reflexivas en docentes en formación. El resultado obtenido es una escala compuesta por 20 ítems que se agrupan en tres factores distintos entre sí, con buenos indicadores de consistencia interna. Los factores obtenidos corresponden a las tres dimensiones teóricas descritas por Larrivee (2008): prácticas de reflexión superficial (3 ítems), prácticas de reflexión pedagógica (8 ítems) y prácticas de reflexión crítica (9 ítems).

La sub-escala que mide la reflexión superficial incluye ítems que refieren a la flexibilidad y al cambio en las prácticas docentes, cuestionando implícitos sobre la enseñanza y el aprendizaje, acogiendo la diversidad en la sala de clases y poniendo su énfasis en resultados de corto plazo.

La sub-escala que mide la reflexión pedagógica incluye prácticas que ponen su foco en el aprendizaje de todos los estudiantes, en la relación entre este aprendizaje y las prácticas del docente y en cómo mejorar estas últimas. Los ítems en esta sub-escala apuntan a tomar en cuenta las conexiones entre los conocimientos previos con los nuevos, considerar el punto de vista de los estudiantes, considerar las consecuencias de las acciones propias en el aprendizaje de los estudiantes, mostrar compromiso con el aprendizaje permanente y la mejora de la práctica, reconocer la complejidad de la dinámica del aula, hacerse autocríticas sobre la propia enseñanza y considerar que la investigación puede contribuir a la enseñanza.

La sub-escala de reflexión crítica incluye prácticas de carácter más bien metacognitivo y relacionadas con la institución escolar en su amplio sentido. Los ítems de esta sub-escala incluyen el considerar incongruencias, supuestos y premisas que subyacen a la práctica docente, analizar la propia forma de pensar, cuestionar los supuestos y las expectativas sobre los estudiantes, identificar y analizar las implicancias éticas, sociales, políticas de las propias acciones y situar el rol de la escuela en un contexto social, histórico y político más amplio.

La presencia de estas tres dimensiones distintas de la escala refuerza la idea de que estas dimensiones no significan unos niveles organizados linealmente y de creciente complejidad, como alguna bibliografía sugiere (Astudillo et al., 2014; Mena & García, 2013; Wunder, 2003). Aunque Larrivee (2008) señala que lo que llama niveles pre-reflexivo, superficial, pedagógico y crítico puede no desarrollarse de manera consecutiva, la interpretación de las tres sub-escalas que se propone en este artículo se orienta más por el trabajo de Collin et al. (2012), según los cuales lo que usualmente se trata como niveles son más bien dimensiones del trabajo docente. Cada una de estas dimensiones debe ser examinada empleando marcos interpretativos enriquecidos, dada la complejidad de la sala de clases (Schön, 1983). Dicho de otro modo, un profesor que desarrolla solo la reflexión crítica podría pasar más bien como un buen analista de la educación, pero no podría examinar con la misma agudeza lo que constituye el centro del trabajo docente: lo que él hace para promover el aprendizaje de los alumnos.

La escala validada consiste en un instrumento de auto-informe. Implica la recopilación de información sobre las propias percepciones de las prácticas de los sujetos. Con ello, el instrumento permite además que los sujetos se representen a sí mismos como deseen o como ellos creen que es más apropiado (Lam & Bengo, 2003; Podsakoff, MacKenzie & Podsakoff, 2012). De este modo, este tipo de datos accede a una dimensión más simbólica

de la práctica reflexiva, diferente de la información obtenida a través de un análisis directo de las prácticas. Ambos enfoques pueden ser complementarios y enriquecer el conocimiento de la experiencia que los profesores y los estudiantes de pedagogía tienen en el desarrollo de prácticas reflexivas.

El instrumento validado aquí busca ser una contribución a la definición y a la medición de la práctica reflexiva en el contexto Hispanoamericano. El instrumento también puede constituir un aporte a las instituciones formadoras de profesores, facilitando el diagnóstico y seguimiento del desarrollo de las prácticas reflexivas en sus estudiantes. Sin embargo, la escala debe ser adaptada para adecuarse a los contextos culturales en los que se aplique y debe ser sometida a nuevas validaciones, con la finalidad de tener un instrumento robusto. Esto contribuirá no solo a comprender mejor el fenómeno en cuestión, sino que permitirá, además, orientar y apoyar los esfuerzos por desarrollar la práctica reflexiva en profesores en formación. Estos esfuerzos han comenzado a ser frecuentes en distintos países de la región, pero difícilmente llegarán a contribuir al mejoramiento de la formación docente si no hay un sólido sustento que permita comprender qué se desea formar y cómo se puede lograr aquello.

REFERENCIAS BIBLIOGRÁFICAS

- Alberca, R., & Frisancho, S. (2011). Percepción de la reflexión docente en un grupo de maestros de una escuela pública de Ayacucho. *Educación*, 20(38), 25-44.
- Astudillo, C., Rivarosa, A., & Ortiz, F. (2014). Reflexión docente y diseño de secuencias didácticas en un contexto de formación de futuros profesores de ciencias naturales. *Perspectiva Educativa*, 53(1), 130-144.
- Beauchamp, C. (2014). Reflection in teacher education: issues emerging from a review of current literature. *Reflective Practice*, 1-19. doi:10.1080/14623943.2014.982525
- Castellanos, S., & Yaya, R. (2013). La reflexión docente y la construcción de conocimiento: una experiencia desde la práctica. *Sinéctica*, 41, 2-18.
- Collin, S., Karsenti, T., & Komis, V. (2012). Reflective practice in initial teacher training: critiques and perspectives. *Reflective Practice*, 14(1), 104-117. doi:10.1080/14623943.2012.732935
- Concha, S., Hernández, C., del Río, F., Romo, F., & Andrade, L. (2013). Reflexión pedagógica en base a casos y dominio de lenguaje académico en estudiantes de cuarto año de pedagogía en educación básica. *Calidad en la educación*, 38, 81-113.
- Consejo Nacional de Educación. (2014). Índices. Matrícula de Educación Superior de pregrado y postgrado 2005-2013 (Base de datos). Recuperado de http://www.cned.cl/public/Secciones/SeccionIndicesPostulantes/Indices_Sistema.aspx
- Escuela Universitaria de Magisterio Universidad Complutense de Madrid. (s/f). Grado de maestro en educación primaria. Competencias generales, transversales y específicas. Madrid: Escuela Universitaria de Magisterio, Universidad Complutense de Madrid.
- Facultad de Educación de la Universidad de Castilla-La Mancha. (2015). Competencias del Título de Grado en Maestro en Educación Infantil. Recuperado de <https://http://www.uclm.es/cr/educacion/gradoEducacionInfantil/competencias.asp>
- Facultad de Educación de la Universidad Complutense de Madrid. (s/f). Competencias. Memoria verificada del grado en maestro en educación infantil Madrid: Facultad de Educación, Universidad Complutense de Madrid.
- Flotts, M. P., & Abarzúa, A. (2011). El modelo de evaluación y los instrumentos. En J. Manzi, R. González y Y. Sun (Eds.), *La evaluación docente en Chile* (pp. 37-61). Santiago de Chile: MIDE

- UC, Centro de Medición. Pontificia Universidad Católica de Chile.
- Fourés, C. (2011). Reflexión docente y metacognición. Una mirada sobre la formación de formadores. *Zona próxima*, 14, 150-159.
- González-Weil, C., Gómez, M., Ahumada, G., Cisternas, A., Pérez, J. L., & Santana, J. (2014). Contribución del trabajo colaborativo en la reflexión docente y en la transformación de las prácticas pedagógicas de profesores deficiencia escolares y universitarios. *Pensamiento Educativo*, 51(2), 75-85.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255-282.
- Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Jaeger, E. (2013). Teacher Reflection: Supports, Barriers, and Results. *Issues in Teacher Education*, 22(1), 89-104.
- Kalk, K., Luik, P., Taimalu, M., & Täht, K. (2014). Validity and reliability of two instruments to measure reflection: a confirmatory study. *Trames*, 18(2), 121-135.
- Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K., . . . Yeung, E. (2000). Development of a Questionnaire to Measure the Level of Reflective Thinking. *Assessment and Evaluation in Higher Education*, 25(4), 381-395. doi:10.1080/026029300449272
- Lam, T. C. M., & Bengo, P. (2003). A Comparison of Three Retrospective Self-reporting Methods of Measuring Change in Instructional Practice. *American Journal of Evaluation*, 24(1), 65-80. doi:10.1177/109821400302400106
- Larrivee, B. (2008). Development of a tool to assess teachers' level of reflective practice. *Reflective Practice: International and Multidisciplinary Perspectives*, 9(3), 341-360.
- Lee, H.-J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching And Teacher Education*, 21(6), 699-715.
- Leijen, Ä., Valtna, K., Leijen, D. A. J., & Pedaste, M. (2012). How to determine the quality of students' reflections? *Studies in Higher Education*, 37(2), 203-217. doi:10.1080/03075079.2010.504814
- Loughran, J. J. (2002). Effective reflective practice - In search of meaning in learning about teaching. *Journal of Teacher Education*, 53(1), 33-43. doi:10.1177/0022487102053001004
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, 4(1), 84.
- Meckes, L. (2014). Estándares y formación docente inicial. En UNESCO/OREALC (Ed.), *Temas críticos para formular nuevas políticas docentes en América Latina y el Caribe: el debate actual* (pp. 53-109). Santiago de Chile: UNESCO/OREALC.
- Mena, J. J., & García, M. L. (2013). El papel de la reflexión crítica en la construcción de conocimiento práctico docente. *Tendencias pedagógicas*, 22, 197-210.
- Ministerio de Educación, Ciencia y Tecnología de Argentina. (2007). *Lineamientos Curriculares Nacionales para la Formación Docente Inicial*. Argentina: Ministerio de Educación, Ciencia y Tecnología de Argentina.
- Ministerio de Educación de Perú. (2012). *Marco del Buen Desempeño Docente*. Lima: Ministerio de Educación de Perú.
- Podsakoff, P., MacKenzie, S., & Podsakoff, N. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539-569.
- Russell, T. (2005). Can reflective practice be taught? *Reflective Practice*, 6(2), 199-204. doi:10.1080/14623940500105833
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8(2), 23-74.
- Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think In Action*. USA: Basic

Books.

- Sobral, D. (2001). Medical students' reflection in learning in relation to approaches to study and academic achievement. *Medical Teacher*, 23(5), 508-513.
- Standal, Ø., & Moe, V. (2013). Reflective practice in physical education and physical education teacher education: A review of the literature since 1995. *Quest*, 65(2), 220-240.
- van Manen, M. (1977). Linking Ways of Knowing with Ways of Being Practical. *Curriculum Inquiry*, 6(3), 205-228.
- Wunder, S. (2003). Preservice teachers' reflections on learning to teach elementary social studies. *Reflective Practice*, 4(2), 193-206.

