



XVI
Congreso Nacional de
Investigación Educativa
CNIE-2021

Validación psicométrica de un cuestionario para evaluar la competencia docente

César Augusto Cardeña Ojeda

Escuela Normal de Ticul
cesarcardenaojeda@gmail.com

Área temática 12. Evaluación educativa.

Línea temática: Evaluación docente y desarrollo profesional.

Tipo de ponencia: Reporte final de investigación.



Resumen

Este trabajo presenta el proceso seguido para validar los atributos métricos de un cuestionario para evaluar la competencia docente de profesores de una escuela normal pública en México, mediante la valoración de los alumnos. Se siguió un método cuantitativo, en el que fue relevante contar con la participación de profesores de la misma institución, pero también, de pares académicos de universidades reconocidas de la región, para fortalecer la tarea de construcción del instrumento, que tuvo la finalidad de sustituir a la versión precedente, en atención a dos intereses específicos: primero, ajustar su contenido a las precisiones del modelo de la educación normal actualizado en 2018, y principalmente, asegurar en el instrumento, una herramienta técnica que evalúe de manera pertinente y con orientación formativa, el desempeño de los docentes de la institución, después de la crisis didáctica enfrentada por la pandemia de la COVID 19. El instrumento construido fue sometido a diversas pruebas estadísticas, en dos etapas generales: una primera, en la que se construyó el modelo preliminar, que se piloteó con una muestra del 70% de la población, lo que generó una robusta base de datos para revisar los atributos métricos del cuestionario, con técnicas de inferencia; en la segunda etapa, se llevó a cabo un segundo piloteo, sensando a la población, y aplicando solamente la revisión de la confiabilidad de las respuestas recabadas, con dos técnicas diferentes. Al término del proceso, se encontraron adecuados valores técnicos que apoyaron el uso del inventario construido en las tareas de evaluación institucional.

Palabras clave: Validación psicométrica; cuestionario; competencia docente.

Introducción

La evaluación del desempeño del profesorado es una actividad clave para apoyar el aseguramiento de la calidad y la mejora de un programa educativo: tiene potencial para impactar en el ejercicio profesional del propio docente, pero también en el del colegio y de la propia institución, en tanto que implica la reflexión de los procesos de enseñanza y los resultados de aprendizaje, de los cuales toda la comunidad escolar son partícipes (Guevara, Meléndez, Ramón, Sánchez, & Tirado, 2016; Instituto Nacional para la Evaluación de la Educación, 2014; Red Iberoamericana de Investigadores de la Evaluación de la Docencia, 2008).

Reconociendo la importancia de la evaluación docente, ésta se constituyó en un tema de principal interés para el área de administración académica de una escuela normal pública de Yucatán, México, a razón de varias situaciones problemáticas:

En principio, al reconocer que los procesos y herramientas empleadas en la evaluación institucional contaban más de 14 años sin haber recibido alguna actualización o adecuación. En esos años, el programa educativo impartido en la escuela ya había afrontado tres cambios o actualizaciones de su plan de estudio: la 'prueba de aula' o piloto 2011, que sustituyó al plan de estudio 1997, el Plan 2012 (Secretaría de Educación Pública, 2012), y el actual Plan 2018 (SEP, 2018);, por tanto, se percibió la necesidad de revisar las actividades e instrumentos de evaluación usados en la institución desde al menos el plan 1997, con la intención de ajustarlos a las características del programa de formación vigente (Luna, & Torquemada, 2008).

También se generó el interés de asegurar la pertinencia técnica de este instrumento, ya que en los archivos del departamento de evaluación institucional, no se contaba con información sobre el proceso seguido para su construcción, y principalmente, porque presentaban evidentes deficiencias estructurales que podían estar causando algún error sistemático en sus mediciones (Kerlinger, & Lee, 2002), como por ejemplo, la repetición de reactivos, doble sentido en el objeto de medida o indicador (Namakforoosh, 2014), una escala de calificaciones reducida, principalmente; al respecto, Cohen (1987) advirtió con claridad que la variabilidad técnica, es decir, la inconsistencia en la calidad o pertinencia de los instrumentos usados para el acopio de datos, anula su validez interna y también su capacidad de generalización.

Adicionalmente, se consolidó el interés de contar con un instrumento válido y confiable que permitiera un adecuado seguimiento formativo al desempeño de los docentes, una vez retomadas las actividades escolares presenciales de la institución, y después de las severas complicaciones didácticas que se enfrentaron a razón de la pandemia por la COVID 19; contar con esta información sería de fundamental importancia para orientar acciones estratégicas que subsanen los perjuicios que el confinamiento ocasionó en el aseguramiento de la formación de los alumnos, desde el mes de marzo de 2020, cuando en el estado de Yucatán se decretó la suspensión de labores escolares presenciales.

De este panorama se planteó la pregunta sobre si se encontraba disponible algún instrumento validado, que se ajustara a las necesidades y características del contexto del programa educativo impartido por la institución, para tramitar lo conducente para su uso y aprovechamiento; sin embargo, al realizar la búsqueda correspondiente en la red libre de internet, no se encontró un caso que satisficiera cabalmente lo requerido, aún entre escuelas normales: en unos casos, el contexto, o el propio objeto de medida presentaba diferencias a lo requerido, y en otros, no se encontraba exhibida la información respecto del proceso de validación técnica de los materiales empleados.

La situación anterior secunda la consideración general de que, en las escuelas normales de México, la evaluación del profesorado aún se constituye en un reto significativo del cual ocuparse (Latapí, 2004; Noriega, Castro, González, & Figueroa, 2018).

Dado lo anterior, se planteó el objetivo de construir y validar los atributos métricos de un instrumento para la valoración del desempeño docente de la institución.

Este objetivo sería uno particular, de un plan más amplio dirigido a mejorar, no solo los materiales de evaluación de la escuela, sino también los procesos para su administración, y el análisis y uso de los datos recabados.

Desarrollo

Para abordar el objetivo planteado, se recurrió a un procedimiento cuantitativo, basado en pruebas estadísticas, principalmente, sobre validez y confiabilidad.

La primera decisión metodológica tomada en este proyecto, fue sobre el diseño general del instrumento: de ello, se optó por usar un inventario que recabara la valoración de los alumnos respecto del desempeño didáctico de sus profesores por cada semestre.

La evaluación del desempeño de los profesores a través de la opinión de los alumnos es la forma más común que ha adoptado esta actividad en las instituciones de educación superior, por sobre otras opciones como la evaluación a través de pares, el portafolio o la autoevaluación (Elizalde, & Reyes, 2008).

Es así que a lo largo de los años, la encuesta de opinión a los estudiantes ha prevalecido como una de las principales estrategias de evaluación docente en las instituciones de educación superior (y en algunos casos, la única), aunque sí se ha transitado de formatos manuales a la incorporación de recursos tecnológicos para su implementación y para el procesamiento de los datos recabados en ella (Cisneros, & Stake, 2010).

Posteriormente, se realizó la revisión de fuentes teóricas para conceptualizar y delimitar el objeto de medida, Competencia docente, el cual quedó establecido para este trabajo como: el desempeño del profesor, que moviliza sus conocimientos, habilidades y actitudes con el fin de generar en sus alumnos experiencias de aprendizaje (Rueda, & Díaz-Barriga, 2000; D'Agostino, & Kosegarten, 2015).

Este concepto se distinguió de la noción de 'Desempeño docente', que puede abarcar rasgos de práctica del profesor, que exceden su labor directa con los estudiantes para la promoción de aprendizajes, es decir, su tarea de enseñanza, y que puede extender su ocupación a otros elementos, como las funciones sustantivas (investigación o tutoría, por ejemplo), la calidad de la interacción con sus compañeros de trabajo (colaboración, relaciones interpersonales, etc.) o su identificación o apego a la institución (Comités Interinstitucionales para la Evaluación de la Educación Superior, 2015; Zabalza, 2003).

Dada esta delimitación conceptual, se insistió en la revisión de trabajos existentes, aproximados o similares, en donde se hubiera realizado la validación técnica de instrumentos para evaluar el desempeño de profesores de programas de educación superior, a fin de encontrar procedimientos y productos específicos que pudieran ser incorporados al instrumento en construcción; de esta revisión se encontraron referentes teóricos e ideas sobre indicadores sobre los cuales elaborar la versión de instrumento específico para la institución; entre los trabajos revisados se encuentran los de Márquez y Madueño (2016), Pimienta (2014) y Vera, Rodríguez, Medina y Gerardo (2012), y con particular énfasis, los casos de trabajos elaborados en escuelas normales del país, como los de Vera, Montoya, y Valdés (s. f) y el de Martínez, Guevara, y Valles, 2016, este último, realizado en un contexto de educación rural.

A partir del análisis acucioso de estos referentes, se redactaron 14 reactivos iniciales, sin ubicarse en indicadores específicos, dada la diversidad de términos y posturas encontradas en los referentes teóricos.

Teniendo este modelo preliminar del instrumento se realizó una sesión de trabajo grupal con siete profesores de la escuela normal, quienes eran al momento, los mejor evaluados en cuanto a su práctica docente, de acuerdo a la información provista por los recursos hasta entonces disponibles: a) el cuestionario de opinión vigente que era administrado a todos los alumnos al término de cada semestre, y b) minutas de las sesiones de trabajo de alumnos representantes de grupo con el subdirector académico de la institución, en las que se trataban diversos temas relacionados con el funcionamiento de la escuela, entre ellos, el desempeño de los docente en la conducción de sus cursos.

Fue importante incorporar a una representación de profesores de la escuela en este trabajo, en atención a la sugerencia de Rueda y Luna (2008) de privilegiar la perspectiva de los profesores en un proceso de evaluación que les atañe tan directamente, como ser los sujetos de evaluación; lo anterior, no únicamente tuvo la finalidad de generar en los docentes la percepción positiva de sentirse involucrados en el trabajo, sino que también tuvo el genuino interés de contar con su conocimiento y experiencia acumulados, para fortalecer los referentes teóricos consultados.

En esta sesión de trabajo los profesores evaluaron los reactivos del modelo preliminar en cuanto a dos cuestiones: a) su validez de contenido, respecto de la definición conceptual construida del objeto de medida, y b) la adecuación técnica de su redacción, que a su vez, se compuso de los siguientes rasgos deseables en cada construcción: claridad, concreción, uni-direccionalidad, vocabulario acorde a la maduración y nivel educativo de

los encuestados, y evitar frases con expresiones extremas (Airasian, 2002; Namakforoosh, 2014; Tornimbeni, Pérez, & Olaz, 2014).

Por cada reactivo, se expusieron argumentaciones para conservarlos íntegros, realizarles modificaciones o eliminarlos, a la vez que se estuvo en la posibilidad de proponer nuevos reactivos, acompañando la sugerencia de su correspondiente argumentación.

Después de tomar en cuenta las observaciones de los profesores de la institución, se envió el instrumento depurado a dos evaluadores externos para que también apoyaran a mejorar el contenido del material con un ejercicio de jueceo ciego; los dos jueces que apoyaron en esta actividad fueron los doctores Pedro Sánchez Escobedo y Francisco Barroso Tanoira, profesores investigadores de programas de educación en la Universidad Autónoma de Yucatán y Universidad Anáhuac Mayab, respectivamente, quienes fueron considerados jueces idóneos para esta tarea, al cubrir ampliamente con los criterios sugeridos para el caso por Skjong y Wentworth (2000): a) experiencia y experticia en la actividad y tema, b) reputación en la comunidad, c) disponibilidad, d) imparcialidad.

En este caso, además de solicitarles una evaluación comentada para cada reactivo, también se solicitó calificarlos con uno de cuatro dictámenes de pertinencia; con los puntajes asignados a los reactivos se realizó una prueba de concordancia entre los dos observadores, empleando la prueba Kappa de Cohen, en la que se encontró un coeficiente de .720 con una significancia estadística de .000, que indicó una adecuada correspondencia entre las mediciones, lo que a su vez corroboró la adecuación del proceso de juicio ciego (Landeró, & González, 2006; Olvera, 2015).

De estos trabajos, los 14 reactivos iniciales se redujeron a 12, que fueron llevados a pilotaje: para ello se administró el instrumento a 268 alumnos de la institución, lo que representó el 70% del total; los alumnos fueron divididos en seis grupos de trabajo a los que previamente se les explicó la finalidad de la actividad, y con quienes se revisó el contenido del instrumento, para asegurar su comprensión sobre cada reactivo.

Con los datos recabados de este ejercicio se realizó la prueba de confiabilidad, usando el estadístico Alfa de Cronbach, del que se obtuvo un coeficiente de .876, considerado como 'adecuado' por los criterios convencionales respecto de la medición de la estabilidad de los resultados de instrumentos puestos a prueba (Cronbach, 2004).

A partir de este resultado positivo, se realizó un Análisis Exploratorio de los factores del instrumento: se recurrió al método de extracción por Componentes Principales, con rotación Varimax (Landeró, & González, 2006); en este proceso se determinó como nivel mínimo aceptable de saturación a valores iguales o superiores a 0.4, en atención al 'Criterio de Stevens' (Martínez, 2005; Hair et al., en Zamora, Monroy, & Chávez, 2009).

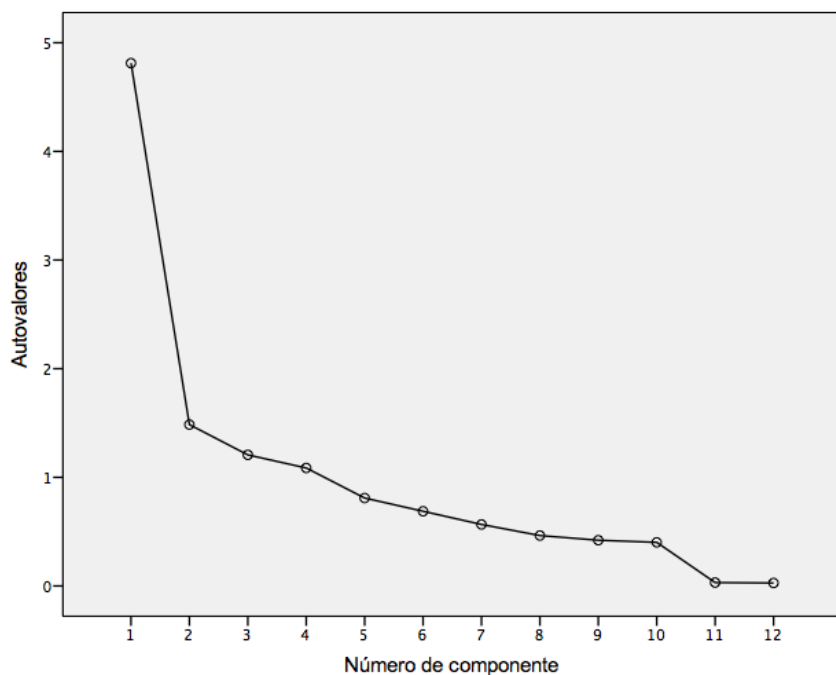
Al conducir este proceso se encontró una varianza total explicada del 62.5% acumulado en tres componentes factoriales principales, como se muestra en la matriz de componentes rotados.

Tabla 1. Matriz de componentes rotados

	Componente		
	1	2	3
Reactivo 8	.827	.411	.213
Reactivo 9	.826	.334	.211
Reactivo 4	.490	.336	.141
Reactivo 6	.475	.012	.147
Reactivo 7	.411	.217	.211
Reactivo 10	.119	.883	.257
Reactivo 12	.131	.826	.284
Reactivo 11	.137	.651	-.064
Reactivo 1	.115	.585	.227
Reactivo 5	.217	.399	.531
Reactivo 2	.247	.143	.525
Reactivo 3	.310	.496	.524

Esta estructura, basada en tres factores, se muestra en la gráfica de sedimentación correspondiente, basada en auto-valores mayores a 1:

Figura 1. Gráfica de sedimentación de los componentes rotados en el Análisis Factorial



Como se puede observar en la figura anterior, el derivado de los componentes señala la distribución usual decreciente de los componentes que se ubican por debajo del auto-valor 1, en tanto que los tres casos por encima de tal valor reflejan la proporción numérica de sus variables puestas en análisis (en este caso, Reactivos).

Por otra parte, al revisar el índice Kaiser –Meyer–Olkin en los datos de este mismo procedimiento, señaló la suficiente adecuación de la muestra al procedimiento de análisis factorial (.786), en tanto que la prueba de

Bartlett corroboró la correlación de identidad entre la variable general (u objeto de medida) y la población involucrada en el proceso de validación técnica del instrumento, al mostrar una significancia estadística inferior al 5% (Zamora, Monroy, & Chávez, 2009).

Con este análisis se pudo estructurar la prueba a partir de tres factores: a) Planeación efectiva, b) Método de gestión del aprendizaje, c) Evaluación formativa; los reactivos correspondientes a estos tres factores se distribuyeron como se muestra:

Tabla 2. Contenido final del instrumento

Indicador	Aspectos
Planeación efectiva.	<ul style="list-style-type: none"> • Muestra previsión en la planeación de su clase (sus materiales, actividades y tareas se encuentran preparadas previamente). • Trabaja en coordinación con otro profesor (al menos una actividad o material usado en su clase se vincula al trabajo de otro profesor). • Emplea recursos didácticos diversificados (los materiales apoyan con claridad el aprendizaje de los temas de estudio). • Incluye Tecnologías de la Información y la Comunicación aplicadas a la educación (las tecnologías gestionan y transforman información que se usa para aprender).
Método de gestión del aprendizaje.	<ul style="list-style-type: none"> • Demuestra dominio disciplinar en los temas que conduce (brinda explicaciones claras y sustentadas en información válida y actualizada). • Estimula el pensamiento crítico (promueve participaciones reflexivas que representan un reto intelectual). • Organiza dinámicas de trabajo acordes a los temas de estudio (las formas de organización del trabajo de aula y extra aula responden a los objetivos de aprendizaje). • Propone tareas con clara aplicación para mi formación profesional (las tareas me ayudan a aprender o a fortalecer lo que he aprendido). • Favorece un clima de trabajo propicio para el aprendizaje (ofrece y promueve actitudes de respeto y colaboración en el grupo).
Evaluación formativa.	<ul style="list-style-type: none"> • Evalúa con sentido formativo (la evaluación me permite monitorear mi aprendizaje y mejorar mis aprendizajes alcanzados). • Evalúa en forma estratégica (usa estrategias y herramientas de evaluación coherentes con los aprendizajes que se pretenden evaluar). • Propicia la reflexión sobre la importancia de la función docente (promueve la autoevaluación sobre mi desempeño competente y ético).

Esta versión final fue turnada al departamento de Evaluación de la institución para su uso a partir del ciclo escolar 2018 – 2019.

Con los datos recabados por el departamento mencionado, se realizó una nueva revisión de la confiabilidad de los datos, encuestando en esta ocasión a 317 alumnos, lo que constituyó un censo poblacional. De nueva cuenta, se usó el modelo Alfa de Cronbach, de lo que se encontró un coeficiente de .820, considerado 'fuerte', según los criterios referidos anteriormente.

Asimismo, se realizó una prueba Test-retest, empleando el coeficiente momento-producto de Pearson, con los datos acopiados de las dos administraciones del inventario (pilotaje y primera aplicación del instrumento

pilotado), encontrándose un coeficiente de relación estadística igual a .868 (sig.= .004), que indica un aceptable nivel de estabilidad sobre el tiempo (Drost, 2011).

Conclusiones

Al término de este proceso, se consideró que el instrumento construido para evaluar el desempeño didáctico del profesor cuenta con adecuados atributos métricos para ofrecer resultados válidos y confiables de su administración, y que se encuentra alineado a los enfoques vigentes de la educación superior y del ámbito de la educación normal.

Por otra parte, la actualización y validación de este instrumento plantea diversos compromisos y retos para la institución:

Por ejemplo, se plantea la oportunidad de ampliar y diversificar las acciones de evaluación que se realizan en la institución, incorporando otras estrategias de alto beneficio individual y colectivo, como la autoevaluación o la coevaluación entre pares (Rizo, 1999), que son actividades que implican el desarrollo de la responsabilidad del profesor por examinar su práctica y la de sus colegas, con la finalidad primordial de mejorarla (Carpintero, González, & Cabezas, 2014; Robalino, & Körner, 2007).

Otro reto es otorgar oportunidades de formación a los responsables del departamento de evaluación, para apropiarse de los conocimientos técnicos de alta especialización que demanda la actividad a su cargo (Rueda, 2008).

Especialmente, un reto mayor está en el uso formativo de la información recabada de la evaluación, para evitar que ésta cree ideaciones desvirtuadas, como la percepción de futilidad, si el proceso de evaluación termina con el levantamiento de los datos y después de esto, no ocurre nada sino hasta el siguiente ejercicio similar; el sentido de utilitarismo, que se origina cuando las evaluaciones se traducen en documentos para engrosar expedientes concursables en las contrataciones o promociones de puesto, o la percepción de ser una herramienta en manos de los directivos para el control laboral de sus subordinados.

Sin duda que el anterior es un desafío que entraña una dificultad sumo compleja, pues se ha documentado que los profesores suelen ofrecer resistencia a estas actividades que –por cierto- ellos mismos deben promover entre sus estudiantes; y esta resistencia ocurre a razón de no otorgarle confianza al proceso, a sus herramientas o a sus implementadores, cuando perciben deficiencias técnico-metodológicas, o porque la evaluación de desempeño se ha asociado precisamente a situaciones no formativas, como por ejemplo, la tendencia a destacar lo negativo, desatender a la diversidad de los profesores, o simplemente, porque existe poca tradición en ella (Guzmán, 2018).

Por lo tanto, el reto más importante por salvar es el de generar en el profesorado un genuino convencimiento de la necesidad de evaluar su práctica, lo que adquiere dimensiones relevantes en la etapa de docencia post-COVID 19, en la que se avecina un trabajo sumamente arduo de reflexión y evaluación, para identificar el nivel de impacto de esta etapa de trabajo en confinamiento, en la formación de los alumnos, y, principalmente, proponer acciones para resarcir dicho impacto.

Sin embargo, los retos y dificultades que se enfrentan en el propósito de consolidar la evaluación institucional no es un argumento suficiente para su renuncia; así, se espera que este aporte motive la revisión de los procesos seguidos en la escuela normal sobre esta tarea de gran importancia para su funcionamiento, y sobre todo, se aboque al uso formativo de sus resultados para el desarrollo profesional de su profesorado y la consecuente mejora del servicio brindado a los estudiantes.

Referencias

- Carpintero M., E., González B., C., & Cabezas G., D. (2014). Evaluación Integral en docentes. Perfiles de docentes con respecto su percepción de la evaluación. *Estudios Pedagógicos*, XL (1),61-74. Recuperado de <http://www.redalyc.org/articulo.oa?id=173531772004>
- Cisneros, E., & Stake, R. (2010). La evaluación de la docencia en educación superior. De evaluaciones basadas en opiniones de estudiantes a modelos por competencias. *Revista Iberoamericana de Evaluación Educativa*, 3 (1), Pp. 219-131. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3693162>
- Cohen, P. A. (1987). *A critical analysis and Reanalysis of the Multisection Validity Meta-Analysis*. Recuperado de https://www.researchgate.net/publication/234572005_A_Critical_Analysis_and_Reanalysis_of_the_Multisection_Validity_Meta-Analysis
- Comités Interinstitucionales para la Evaluación de la Educación Superior. (2015). *Guía para la Autoevaluación de Programas de Educación Superior*. México: CIEES.
- Cronbach, L. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64 (3). Recuperado de http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.575.7830yrep=rep1yt_type=pdf
- D'Agostino, S, & Kosegarten, J. (2015). Reevaluating Teaching Evaluations. *Liberal Education*, v101 n3 Sum 2015. ERIC Number: EJ1095212
- Drost, E. A. (2011). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, 38 (1). Pp105-123.
- Elizalde, L., & Reyes, R. (2008). Elementos clave para la evaluación del desempeño de los docentes. *Revista Electrónica de Investigación Educativa* (Número Especial). Recuperado de <http://redie.uabc.mx/NumEsp1/contenido-elizaldereyes.html>
- Guevara, G., Meléndez, M. T., Ramón, F. E., Sánchez, H., & Tirado, F. (2016). *La evaluación docente en México*: México: Fondo de Cultura Económica.

- Guzmán, F. (2018). La experiencia de la evaluación docente en México: análisis crítico de la imposición del Servicio Profesional Docente. *Revista Iberoamericana de Evaluación Educativa*, 11(1), 135-158. Recuperado de <https://doi.org/10.15366/riee2018.11.1.008>
- Instituto Nacional para la Evaluación de la Educación. (2014). *La evaluación de los maestros*. México: INEE. Recuperado de https://publicaciones.inee.edu.mx/buscadorPub//P1/C/231/P1C231_10E10.pdf
- Kerlinger, F., & Lee, H. (2002). *Investigación del comportamiento. Métodos de investigación en ciencias sociales*. México: McGraw Hill.
- Landero, R., & González, M. T. (2006). *Estadística con SPSS y metodología de la investigación*. México: Trillas.
- Latapí, P. (2004). La política educativa del Estado mexicano desde 2002. *Revista Electrónica de Investigación Educativa*, 6 (2). Recuperado de <http://www.scielo.org.mx/pdf/redie/v6n2/v6n2a11.pdf>
- Luna, E., & Torquemada, A. D. (2008). Los cuestionarios de evaluación de la docencia por los alumnos: balance y perspectivas de su agenda. *Revista Electrónica de Investigación Educativa*, (número especial), pp. 1-15. Recuperado de <http://redie.uabc.mx/redie/article/view/201>
- Márquez, L., & Madueño, M. L. (2016). Propiedades psicométricas de un instrumento para apoyar el proceso de evaluación del docente universitario. *Revista Electrónica de Investigación Educativa*, 18(2), 53-61. Recuperado de <http://redie.uabc.mx/redie/article/view/793>
- Martínez, A. R. (2005). *Psicometría: Teoría de los tests psicológicos y educativos*. España: Síntesis Psicología.
- Martínez, G., Guevara, A., & Valles, M. (2016). El desempeño docente y la calidad educativa. *Ra Ximhai*, 12 (6), 123-134. Recuperado de <http://www.redalyc.org/pdf/461/46148194007.pdf>
- Namakforoosh, M. (2014). *Metodología de la investigación*. 2a Edición. México: Limusa.
- Noriega, J. A., Castro, G, González, N. G., & Figueroa, F. L. (2018). Modelo de autoevaluación y heteroevaluación de la práctica docente en Escuelas Normales. *Educação e Pesquisa*, 44. Recueprado de http://www.scielo.br/scielo.php?pid=S151797022018000100461yscript=sci_abstractytlng=es
- Olvera, M. C. (2015). *Coeficiente Kappa promedio: un nuevo parámetro para evaluar y comparar el rendimiento de test diagnósticos binarios*. Tesis doctoral. España: Universidad de Granada.
- Pimienta, J. H. (2014). Elaboración y validación de un instrumento para la medición del desempeño docente basado en competencias. *Revista de docencia universitaria*. 12(2) , 231-250.
- Red Iberoamericana de Investigadores de la Evaluación de la Docencia. (2008). Reflexiones sobre el diseño y puesta en marcha de programas de evaluación de la docencia. *Perfiles Educativos*, XXX (122), 136-140.
- Rizo, H. (1999). Evaluación del docente universitario. *Revista Electrónica Interuniversitaria de Formación del Profesorado*. 2 (1). Recuperado de https://www.aufop.com/aufop/uploaded_files/articulos/1224341864.pdf
- Robalino, M., & Körner, A. (coords). (2007). *Evaluación del desempeño y carrera profesional docente. Un estudio comparado entre 50 países de América y Europa*. Santiago de Chile: OREALC/UNESCO
- Rueda, M. (2008). La evaluación del desempeño docente en la universidad. *Revista Electrónica de Investigación Educativa* (Número Especial). Recuperado de <http://redie.uabc.mx/NumEsp1/contenido-rueda.html>
- Rueda, M., & Díaz-Barriga, F. (Compiladores). (2000). *Evaluación de la docencia. Perspectivas actuales*. México: Paidós Educador.
- Rueda, M., & Luna, E. (2008). La docencia universitaria y su evaluación. *Revista Electrónica de Investigación Educativa* (Número Especial). Recuperado de <http://redie.uabc.mx/NumEsp1/contenido-np2008.html>

- Secretaría de Educación Pública. (2012). *Acuerdo número 649 por el que se establece el Plan de Estudios para la Formación de Maestros de Educación Primaria*. México: DOF. Recuperado de <http://www.sep.gob.mx/work/models/sep1/Resource/9721849d-666e-48b7-8433-0eec1247f1ab/a649.pdf>
- Secretaría de Educación Pública. (2018). *Acuerdo número 14/07/18 por el que se establecen los planes y programas de estudio de las licenciaturas para la formación de maestros de educación básica que se indican*. México: DOF. Recuperado de https://www.dof.gob.mx/nota_detalle.php?codigo=5533902&fecha=03/08/2018
- Skjong, R., & Wentworth, B. H. (2000). *Expert Judgement and risk perception*. Recuperado de <http://research.dnv.com/skj/Papers/SkjWen.pdf>
- Tornimbeni, S., Pérez, E., & Olaz, F. (2014). *Introducción a la psicometría*. Buenos Aires: Paidós.
- Vera, J. A., Montoya, G., & Valdés, A. A. (s. f.). *Perfiles de docentes en escuelas normales del estado de Sonora*. Recuperado de <https://www.ciad.mx/archivos/desarrollo/publicaciones/PUBLICACIONES/Produccion%20Academica/Articulos/2014/205.pdf>
- Zabalza, M. A. (2003). *Competencias docentes del profesorado universitario*. España: Narcea.
- Zamora, S., Monroy, L., & Chávez, C. (2009). *Análisis factorial: una técnica para evaluar la dimensionalidad de las pruebas*. México: Centro Nacional de Evaluación para la Educación Superior, A. C.