AÑO 3. Nº 05 FEBRERO - JULIO DE 2019
ISSN: 2594-0759

VALIDACIÓN DE UN INSTRUMENTO DE MEDICIÓN PARA EL DISEÑO DE INSTRUMENTOS DE EVALUACIÓN A TRAVÉS DEL JUICIO DE EXPERTOS

VALIDATION OF A MEASUREMENT INSTRUMENT FOR THE DESIGN OF EVALUATION INSTRUMENTS THROUGH EXPERT JUDGMENT

Lic. *Roberto Carlos Vea-Solano.*
Escuela Normal Experimental de El Fuerte.
rcvs_2702@hotmail.com

Lic. *Roberto Carlos Vea-Solano*

RESUMEN

En la Escuela Normal Experimental de El Fuerte encontramos la falta de instrumentos que permitan evaluar la enseñanza y el diseño de instrumentos de evaluación pertinentes a los retos de la sociedad del conocimiento y que apliquen la socioformación. Por ello, fue necesario realizar la validación de contenido, mediante juicio de expertos, de un instrumento que permitiera generar estrategias para el diseño e implementación de instrumentos de evaluación en alumnos normalistas. Se realizó un estudio instrumental con base en los siguientes pasos: 1) construcción de un instrumento de validación del diseño de instrumentos de evaluación por parte de estudiantes normalistas mediante una estructura de rúbrica socioformativa; y 2) validación de contenido por parte de un grupo de 15 jueces expertos en el área. Una vez concluido el juicio se procesaron los resultados a través del coeficiente de la V de Aiken, arrojando significación superior a los 0.85 tanto en la pertinencia como en redacción de todos los ítems. Lo anterior permitió presentar un instrumento pertinente para evaluar el diseño de instrumentos de evaluación de los estudiantes normalistas, estableciéndose a la vez, la necesidad de estudios complementarios de validez y confiabilidad referentes al tema.

Palabras claves: Evaluación, Instrumentos de medición, Investigación, Juicio

ABSTRACT

In the Normal Experimental School of El Fuerte we find the lack of instruments that allow evaluating the teaching and design of assessment instruments relevant to the challenges of the knowledge society and that apply socioformation. For this reason, it was necessary to validate the content, through expert judgment, of an instrument that would allow the generation of strategies for the design and implementation of evaluation instruments in student teachers. An instrumental study was carried out based on the following steps: 1) construction of a validation instrument for the design of evaluation instruments by student teachers through a socioformative rubric structure and, 2) validation of content by a group of students. 15 expert judges in the area. Once the trial was concluded, the results were processed through the coefficient of the V of Aiken, yielding significance higher than 0.85 both in the relevance and in the writing of all the items; This allowed to present a pertinent instrument to evaluate the design of assessment instruments for normal students, establishing at the same time, the need for complementary studies of validity and reliability related to the subject.

Key Words: *Evaluation, Measurement instruments, Investigation, Judgment*

Lic. *Roberto Carlos Vea-Solano*

INTRODUCCIÓN

El proceso evaluativo en el contexto educativo se ha constituido en un aspecto central de la etapa enseñanza/aprendizaje, por lo que los instrumentos que se utilicen deben ser pensados y diseñados detenidamente, con la finalidad de permitir una igualdad de oportunidades a los estudiantes evitando evaluaciones estandarizadas. Las Escuelas normales han tenido desde antaño la gran responsabilidad de formar a los mejores profesores del país, es por ello que la Normal Experimental de El Fuerte ha asumido el compromiso con la calidad de la educación que busca el sistema educativo en nuestro país y que tanto añora la sociedad mexicana.

La implementación de una metodología para el diseño de instrumentos de evaluación es indiscutiblemente necesaria, por los tiempos en los que nos encontramos al vivir en una sociedad del conocimiento en la que constantemente estamos enfrentando cambios por lo que debemos evaluar el desempeño de nuestras habilidades y competencias adquiridas a lo largo de nuestra formación. Para que este proceso de evaluación pueda generarse en la mayor de las igualdades posibles de acuerdo a las características propias de cada contexto y cada estudiante, es necesario la utilización de instrumentos evaluativos que ayuden a cumplir con este objetivo. Instrumentos de evaluación definidos por (Ibarra y Rodríguez, 2011: 71-72) como “herramientas reales y tangibles utilizadas por la persona que evalúa para sistematizar sus valoraciones sobre los diferentes aspectos”. Algunos ejemplos son: las listas de control, las escalas de estimación, las rúbricas, las escalas de diferencial semántico, las matrices de decisión o incluso instrumentos mixtos donde se mezclen más de uno.

A pesar de la existencia de autores que señalan la importancia de los instrumentos de evaluación en los procesos educativos y de la existencia de algunos estudios que validan el diseño de instrumentos de evaluación de manera individual, son pocos los trabajos que validan una metodología de diseño de estos instrumentos en una Escuela Normal, por lo que es necesario que acciones como estas se llevan a cabo para dar cumplimiento a la educación de calidad que tanto se ha buscado a través del tiempo y de las distintas reformas al sistema educativo de nuestro país, dejando en claro la importancia del presente estudio de validación.

El diseño de una metodología para conocer los instrumentos de evaluación que permitan al futuro docente mejorar su desempeño profesional se constituye en el propósito de este estudio documental, partiendo de los siete los pilares sobre los que descansa la evaluación en la educación superior: ¿por qué evaluar?; ¿cómo evaluar?; ¿qué evaluar?; ¿cuándo evaluar?; ¿quién evalúa?; ¿con qué calidad se evalúa?; ¿qué hacer tras la evaluación? Falchivok, (2005). Es por lo anterior que la metodología diseñada fue evaluada y validada a través de un juicio de 15 expertos en el área educativa.

> **“Evaluar es atribuir valor a las cosas, es afirmar algo sobre su mérito”**
>
> ***Santos Guerra, (1993)***

Considerando lo anterior, los objetivos de esta investigación fueron: 1) Diseñar un instrumento pertinente y práctico para evaluar una metodología para el diseño de instrumentos de evaluación considerando los retos de la sociedad del conocimiento y el enfoque socioformativo; 2) Realizar la validación de contenido del instrumento con un grupo de 15 jueces, para determinar su grado de pertinencia respecto al constructo evaluado y la redacción de las preguntas y opciones de respuesta.

Fundamentación teórica.

El desarrollo de un modelo de educación basado en competencias establece un cambio en el rol que la educación superior tiene en la formación de profesionales que puedan enfrentar los nuevos desafíos y necesidades de la sociedad, junto con nuevas propuestas educativas que asignan un papel activo al estudiante en su proceso de aprendizaje, y que apuntan a un tipo de aprendizaje profundo y significativo, así como, al logro de ciertas habilidades específicas y evaluables. Rosenbluth, Cruzat y Ugarte, (2015).

En la actualidad, en general, la evaluación del aprendizaje en una carrera o programa se realiza esencialmente en las asignaturas y otras unidades de enseñanza aprendizaje complementarias. Tanto los programas de asignatura como los procedimientos de evaluación son usualmente establecidos por los especialistas de cada materia con bastante independencia. Siguiendo a Sanmartí, (2007), consideraremos la evaluación como un proceso basado en recoger información, sea por medio de instrumentos escritos o no escritos; analizar esa información y emitir un juicio sobre ella, tomando decisiones de acuerdo con el juicio emitido. Para (Santos Guerra, 1993: 63-64) “evaluar es atribuir valor a las cosas, es afirmar algo sobre su mérito”.

La evaluación en los nuevos planes resulta uno de los elementos clave que asegura el verdadero cambio respecto la docencia de planes anteriores. Evaluar las competencias, en vez de (o a la vez que) evaluar los contenidos, merece un cambio en la mirada de los profesionales implicados. Para que este proceso se genere con equidad deben existir una variedad de instrumentos que permitan la evaluación; estos son definidos por Hamodi, López y López, (2015) como: “herramientas que tanto el profesorado como el alumnado utilizan para plasmar de manera organizada la información recogida mediante una determinada técnica de evaluación”. Calderón, Gustems, Brown y Pickford, (2013) señalan que la propuesta de rúbricas para regular la evaluación es una auténtica aportación que enriquece la calificación de contenidos conceptuales expresada mediante un simple número.

La inclusión de calificativos de tipo cualitativo junto a evidencias y datos cuantitativos permite mayor eficacia en la evaluación a la vez que informa al alumno de su nivel de formación, no sólo del nivel de conocimientos adquiridos. Por su parte, Casanova, (1998) propone “un modelo evaluador y su metodología” en el que incluye: 1) técnicas para la recogida de datos (observación, entrevista, encuesta, sociometría, coloquio y trabajos del alumnado); 2) técnicas para el

análisis de datos (triangulación y análisis de contenido); y 3) instrumentos para la recogida y el análisis de datos (anecdotario, lista de control, escala de valoración, cuestionario, sociograma y psicograma, diario y grabación). Salinas, (2002) dedica un capítulo de su libro a "instrumentos para tratar de evaluar", donde alude a los exámenes, trabajos, cuadernos, preguntas de clase, exámenes y controles, pruebas objetivas y observación sistemática.

Las propuestas anteriores se ven reflejadas en el Acuerdo 649 por el que se establece el Plan de Estudios para la Formación de Maestros de Educación Básica S.E.P., (2012) de nuestro país y en específico en las diversas asignaturas que conforman este plan que señala la importancia de utilizar las propias tareas de aprendizaje como evidencia, ya que permiten una evaluación del proceso de aprendizaje y no solo de los resultados, para ello el docente debe seleccionar métodos y estrategias acordes para el tipo de desempeño a evaluar, por lo que es posible, utilizar entrevistas, debates, observación del desempeño, proyectos, casos, problemas , exámenes y portafolios, entre otros.

Hay que abogar por un currículum que incluya todo lo importante para que la población disponga de una formación básica integral, que favorezca aprendizajes transferibles, que dote de competencias suficientes como para desenvolverse satisfactoriamente en la vida, que promueva el entendimiento entre los diferentes (al fin, entre todos: cada persona es única e irrepetible), que forme para la propia toma de decisiones…, en fin, que resulte decisivo para conformar un perfil de "especialista en ser persona (como demanda la vida) y de especialista en ser profesional (como demanda la sociedad)" Casanova, (2012).

Para ello, es necesario que los futuros docentes de educación básica reconozcan la importancia de conocer, diseñar y aplicar una diversidad de instrumentos evaluativos acordes con las características de cada evidencia que se pretende que los alumnos elaboren, pero también tomando en cuenta las características propias de la diversidad de alumnos que conformen su aula. Es entonces necesario que los docentes normalistas guíen a los futuros docentes proveyéndoles de insumos de calidad y que permitan que las prácticas profesionales de estos sean exitosas con evidencias de calidad que cumplan con las exigencias del sistema educativo; pero, sin dejar de lado el sentido humano que caracteriza a la profesión docente.

Tal como lo señalan Muñiz y Fonseca, (2008) la importancia de unos instrumentos de evaluación adecuados radica en la trascendencia de las decisiones y las consecuencias que a partir de ellos se derivan, tanto personales como sociales. La eficacia de la evaluación depende de la pertinencia de la combinación de diferentes instrumentos, de la oportunidad en que se administra y de la inteligencia y propiedad del análisis e interpretación de sus resultados De Camilloni, (2000).

Actualmente existen estudios de validación de algunos instrumentos de evaluación pero de manera individual. Arribas, (2004) describe la metodología que debe utilizarse para la construcción y validación de cuestionarios, como escalas de medida que permiten la obtención de datos y su cuantificación con la finalidad última de poder comparar información. Por su parte, Camargo y Pardo, (2008) realizan otro instrumento en el que evalúan siete competencias básicas para el ejercicio de la función docente: planificación curricular, utilización adecuada de diseño metodológico y organización de actividades de enseñanza, competencia científica tecnológica, interacción adecuada con estudiantes, competencia para evaluar, competencia para realizar tutorías, autorreflexión sobre la práctica docente.

Como indican Cashin, (1990) y Ruiz, (2005) citados por Cisneros, Jorquera & Aguilar, (2012), numerosos estudios han examinado la validez de las evaluaciones que los estudiantes hacen de sus maestros y han encontrado que los estudiantes son una de las más importantes fuentes de información sobre la docencia. El Consejo de Evaluación, Acreditación y Certificación de la Calidad de la Educación Superior (CONEAU), (2012) aborda la validez de instrumentos de evaluación, desde la validación externa donde se describe el procedimiento de la validez de contenido por juicio de expertos, las evidencias basadas en el contenido con ejemplos de procedimiento de validez según Gregory y la V de AIKEN.

Metodología

Tipo de Estudio

Se realizó un estudio instrumental de validez de contenido, tipo de estudio que consiste en todos los estudios encaminados al desarrollo de pruebas y aparatos, incluyendo tanto el diseño como la adaptación de estos Montero y León, (2002); Ato, López, y Benavente, (2013). Lo anterior, a través del instrumento denominado "Rúbrica socioformativa para el diseño de instrumentos de evaluación en estudiantes normalistas", integrado por siete dimensiones, con cuatro aspectos a evaluar cada uno como posibles respuestas, siguiendo los niveles de desempeño desde la socioformación diseñados por Tobón, (2014), relacionados con la importancia de la enseñanza y aprendizaje del diseño de instrumentos de evaluación por estudiantes normalistas.

Procedimiento

El estudio de validez de contenido, de acuerdo con Arribas, (2004) se refiere a si el cuestionario elaborado, y por lo tanto los ítems elegidos, son indicadores de lo que se pretende medir; es por ello que este estudio se llevó a cabo mediante las siguientes fases: Primero, se elaboró una rúbrica para el aprendizaje y diseño de instrumentos de evaluación para los estudiantes normalistas, la cual estaba basada en la socioformación, enfoque que forma parte de los enfoques que buscan formar para la sociedad del conocimiento Tobón, (2013) en el contexto de docentes y trabajadores de la educación, con preguntas planteadas de manera, ordenada, objetiva, concreta y precisa, considerando la importancia del instrumento, por ser una herramienta muy valiosa que representa el primer recurso de la investigación García Córdoba, (2004); citado en Dorantes, Hernández y Tobón, (2016).

Se eligió el método juicio de expertos, método de validación útil para verificar la fiabilidad de una investigación que se define como "una opinión informada de personas con trayectoria en el tema, que son reconocidas por otros como expertos cualificados en éste, y que pueden dar información, evidencia, juicios y valoraciones" (Escobar-Pérez y Cuervo-Martínez, 2008: 29). Lo anterior, para la validación de dicha rúbrica, con expertos que fueron elegidos por su nivel de conocimiento y experiencia en el área educativa; para el número de jueces ideal en Cabero y Llorente, (2013), la selección del número de expertos depende de aspectos como la facilidad para acceder a ellos o la posibilidad de conocer expertos suficientes sobre la temática objeto de la investigación. Por otra parte, autores como (Escobar-Pérez y Cuervo-Martínez, 2008: 29) señalan que "el número de jueces que se debe emplear en un juicio depende del nivel de experticia y de la diversidad del conocimiento; mientras que, Escurra, (1988) afirma que el número de jueces puede variar según el instrumento.

Para este juicio se eligieron quince jueces, a los cuales se les enviaron correos electrónicos en donde se les solicitaba su participación y posteriormente se les explicó tanto los indicadores como cada uno de los ítems que conformaban la rúbrica socioformativa definida como instrumentos de evaluación que permiten determinar el nivel de logro o desempeño en la resolución de problemas del contexto, relacionando una serie de indicadores con descriptores de niveles de dominio Tobón, (2013), donde cada indicador contenía una serie de descriptores, el experto debería leer cada descriptor desde el nivel receptivo hasta el estratégico para comprender la secuencia.

El nivel receptivo responde a un nivel insuficiente de aprendizaje, mientras que el nivel estratégico se relaciona con un nivel de alta apropiación del indicador y aprendizaje. El nivel resolutivo responde a los elementos mínimos esperados y se considera básico. El nivel autónomo responde a los elementos de un desempeño independiente, hay argumentación y se resuelven diversos problemas; mientras que en el nivel estratégico se proponen alternativas de cambio con inventiva y originalidad. La forma de calificación o evaluación fue mediante la selección de uno de los niveles que se señalan en cada una de los ítems so-

bre la pertinencia del indicador para la problemática y sobre la claridad de la redacción de cada una de las dimensiones.

Con los resultados arrojados tras las respuestas de los jueces se procedió al cálculo de la concordancia entre estos por medio de la V de Aiken, "técnica que consiste en la medición con precisión de la validez de contenido con un carácter cualitativo de la pertinencia y adecuación de los instrumentos, a través del juicio de expertos" Sáez, (2017). Por último, se elaboraron conclusiones del juicio de expertos y se dieron a conocer datos sobre su posterior utilización.

El instrumento anterior se complementó con un: "Cuestionario de Factores Sociodemográfico", que se incluyó en el formulario de juicio de expertos que tiene como meta recoger datos sociodemográficos de los participantes, respecto a aspectos tales como edad, sexo, nivel educativo, experiencia profesional e investigativa, tal como se señala en la Tabla 1.

Tabla 1.
Datos Sociodemográficos de los Expertos

Expertos		
16 jueces	Sexo	72.6 % hombres 27.4 % mujeres
	Último nivel de estudio:	Licenciatura: 100%
	Áreas de experiencia profesional:	66.7 % Docencia, 6.7 % procesos psicológicos 6.7 % procesos organizacionales 20% Otros
	Número de años de experiencia profesional:	20 % 0 años 6.7 % 1 año 6.7 % 2 años 13.3% 4 años 6.7 % 5 años 20 % 6 años 6.7 % 7 años 6.7 % 8 años 6.7 % 9 años 6.7 % 11 años
	Número de años de experiencia docente:	73.3 % 0 años 13.3 % 1 año 6.7 % 4 años 6.7 % 6 años
	Número de artículos publicados en el área:	80 % 0 13.3 % 1 6.7 % 10
	Número de libros publicados en el área:	93.3 % 0 6.7 % 6
	Número de capítulos de libros publicados en el área:	93.3 % 0 6.7 % 32
	Experiencia en la revisión, diseño y/o validación de un determinado instrumento de investigación:	Si: 46.7 % No: 53.3 %

La rúbrica socioformativa como instrumento para este proceso de validación está compuesta por las siguientes 7 dimensiones: 1) diagnosticar; 2) integrar un equipo de trabajo; 3) revisión de la malla curricular; 4) planteamiento de problemas de contexto; 5) establecimiento de estrategias para mejorar los conocimientos; 6) resolución del problema y 7) evaluación; con cuatro ítems cada una refiriéndose a la pertinencia y a la claridad de cada dimensión, lo que esta descrito de manera sintética en la Tabla 2.

Tabla 2
Dimensiones y Preguntas del Instrumento

DIMENSIONES	ÍTEMS DEL INSTRUMENTO
Dimensión 2. Integrar un equipo de trabajo	Se deberá elegir un alumno de cada grupo para que formen parte de este equipo y obtener información, desde su perspectiva, sobre qué es lo que hace falta para mejorar en el diseño de los instrumentos de evaluación.
Dimensión 3. Revisión de la malla curricular	Con apoyo del equipo integrado se analizarán los cursos que conforman la malla curricular de la Licenciatura y rescatar los instrumentos de evaluación que se sugieren utilizar en el transcurso de cada uno.
Dimensión 4. Planteamiento de problemas de contexto	Se deberán analizar los factores que influyen y los problemas que pueden generarse con la implementación de instrumentos de evaluación inadecuados en las prácticas docentes o bien, la no implementación de estos.
Dimensión 5. Establecimiento de estrategias para mejorar los conocimientos	Se brindará una conferencia-curso o taller (dependiendo de las posibilidades escolares) a un número determinado de estudiantes normalistas para que reconozcan la importancia de los instrumentos de evaluación y la relevancia de estos en el proceso educativo.

Dimensión 6. Resolución del problema	Se elegirá un muestreo de estudiantes normalistas (algunos que estuvieron en el curso-taller y otros que no lo presenciaron) para que elaboren una planeación didáctica donde establezcan la aplicación de instrumentos de evaluación.
Dimensión 7. Evaluación.	Se compararán las planeaciones de quienes estuvieron en la conferencia-curso o taller con la de quienes no fueron parte de él, para evaluar su desempeño y demostrar si esa estrategia permitió una mejora en el uso adecuado de los instrumentos de evaluación

Estudio de la validez de contenido

La validez de contenido se refiere a si el cuestionario elaborado, y por tanto los ítems elegidos, son indicadores de lo que se pretende medir. Se trata de someter e cuestionario a la valoración de investigadores y expertos, que deben juzgar la capacidad de este para evaluar todas las dimensiones que deseamos medir (Arribas, 2004 p.27). A través de la validez de contenido se trata de determinar hasta dónde los ítems de una prueba son representativos de los indicadores del criterio a evaluar de las competencias profesionales que se desea medir, establece la relación del instrumento con los criterios de los elementos de la competencia que pretende medir (CONEAU, 2012).

Como se ha mencionado en líneas anteriores esta validez se realizó mediante la evaluación por parte de 15 jueces, entre ellos: profesionales del área educativa con experiencia de 1 a 8 años en esta área y algunos más estudiantes normalistas; 9 mujeres y 6 hombres, con edades que oscilan entre los 20 y 35 años. La participación fue a través de una plataforma virtual habilitada para colocar el nivel de pertinencia y de redacción en cada uno de los ítems. Para la realización de este estudio se empleó la V de Aiken, que de acuerdo a Fernández, Cayssials y Pérez (2009) "es la verificación de que la muestra de ítems incluida en el test cubra, efectivamente, todos los aspectos o dimensiones relevantes de la variable en estudio a ser medida".

Resultados

Los resultados obtenidos en cuanto al nivel de pertinencia y redacción asignado por los expertos para cada uno de los ítems de la rúbrica socioformativa para el diseño de instrumentos de evaluación por alumnos normalistas, mediante la técnica V de Aiken, indican que el nivel es significativo pues se observa que excede el valor crítico de 0.50 ya que todos los ítems alcanzaron un promedio en la pertinencia ≥ 8.7 y un promedio en su redacción ≥ 8.4 (Tablas 3 y 4) alcanzando ambas medidas valores óptimos (Bulger y Housner, 2007) (Citados por Blázquez 2011); demostrando que la prueba puede ser utilizada para los fines para los cuales fue diseñada (Escobar-Pérez y Cuervo Martínez, 2008), porque se estima que hay concordancia entre los jueces y demuestra que los indicadores son adecuados, aunque susceptibles de mejora, ya que el análisis determina relaciones significativas entre criterios e indicadores, lo que le otorga validez y confiabilidad al instrumento.

Tabla 3.
Resultados de pertinencia en base a la V de Aiken

ÍTEM	PERTINENCIA	P
Ítem 1	0.87	p<0,05
Ítem 2	0.93	p<0,05
Ítem 3	0.89	p<0,05
Ítem 4	0.89	p<0,05
Ítem 5	0.91	p<0,05
Ítem 6	0.93	p<0,05
Ítem 7	0.93	p<0,05

Tabla 4.
Resultados de redacción en base a la V de Aiken

ÍTEM	PERTINENCIA	P
Ítem 1	0.84	p<0,05
Ítem 2	0.91	p<0,05
Ítem 3	0.87	p<0,05
Ítem 4	0.89	p<0,05
Ítem 5	0.91	p<0,05
Ítem 6	0.93	p<0,05
Ítem 7	0.91	p<0,05

Lo anterior, tiene estrecha relación y da argumentos necesarios para demostrar la importancia del presente trabajo en base a los resultados arrojados por la V de Aiken después del juicio de expertos realizado para la verificación de la pertinencia y redacción de los ítems (Tabla 5) siendo estos valores considerados como muy adecuados Pendfield y Giacobbi, (2004).

Tabla 5.
Resultados de pertinencia y redacción en base a los resultados de la V de Aiken

ÍTEM	PERTINENCIA	REDACCIÓN
Ítem 1	0.87	0.84
Ítem 2	0.93	0.91
Ítem 3	0.89	0.87
Ítem 4	0.89	0.89
Ítem 5	0.91	0.91
Ítem 6	0.93	0.93
Ítem 7	0.93	0.91
Total	0.90	0.89

Discusión

A partir del estudio llevado a cabo, puede señalarse que la evaluación es un proceso que requiere de la recolección de información respecto del objeto que se está evaluando. En materia educativa, aquella recolección se realiza principalmente a través de instrumentos de evaluación, que pueden ser definidos como todo aquello que permite obtener información respecto a la adquisición y grado de logro de un aprendizaje de los estudiantes Castillo, (2003); Pimienta, (2008). Además, que los instrumentos de evaluación son las herramientas que tanto el profesorado como el alumnado utilizan para plasmar de manera organizada la información recogida mediante una determinada técnica de evaluación, Hamodi y López, (2015); o bien, "herramientas reales y tangibles utilizadas por la persona que evalúa para sistematizar sus valoraciones sobre los diferentes aspectos" Ibarra y Rodríguez, (2011).

Es importante rescatar que la evaluación mediante el juicio de expertos, método de validación cada vez más utilizado en la investigación, consiste, básicamente, en solicitar a una serie de personas la demanda de un juicio hacia un objeto, un instrumento, un material de enseñanza, o su opinión respecto a un aspecto concreto (Cabero y Llorente, 2013:14). Se trata de una" técnica cuya realización adecuada desde un punto de vista metodológico constituye a veces el único indicador de validez de contenido del instrumento de recogida de datos o de información" Escobar Pérez, (2008). En donde validez y fiabilidad son los dos criterios de calidad que debe reunir todo instrumento de medición tras ser sometido a la consulta y al juicio de expertos con el objeto de que los investigadores puedan utilizarlo en sus estudios. Hamodi y López, (2015).

Existe una diversidad de juicios de expertos sobre distintas temáticas, incluso, algunos hablan de instrumentos de evaluación; pero, abordándolos de manera individual, no valorando una metodología que permita reconocer las deficiencias y los obstáculos a los que los futuros docentes se han enfrentado en su aprendizaje y, más aún en su diseño e implementación en el aula de educación básica, espacios donde realizan sus prácticas profesionales en donde se desarrollarán la mayor parte de su tiempo.

Diversos estudios señalan la importancia del juicio de expertos para asegurar la pertinencia de los instrumentos, ya que con mucha frecuencia se diseñan instrumentos con bajo grado de coherencia con la teoría y el entorno. De esta manera, se puede plantear que el instrumento validado está acorde con los avances teóricos y puede ser relevante en futuras investigaciones. Casanova, (2012) señala que las características de la sociedad actual han derivado en una realidad pluricultural (que tiende a la interculturalidad con fuerza), en gobiernos democráticos que deben respetar las diferencias de sus ciudadanos y enriquecerse con ellas, en la disponibilidad de múltiples medios tecnológicos que permiten el acceso a la información de diferentes sectores sociales, etc. Para la sociedad se ha convertido en una necesidad rele-

vante que la educación superior evidencie una concepción renovada, que denote una práctica moderna, flexible y centrada en el mejoramiento de la competencia de los profesores la cual, con el apoyo de políticas oficiales e institucionales, busque y propenda por el mejoramiento de la competencia académica de los profesores Díaz, (2000).

Si hablamos de diferencia en los ciudadanos y cambios para la mejora de la educación, es importante entonces que esas diferencias sean reconocidas desde los procesos educativos y, muy en especial, en los momentos de evaluación. Siguiendo a Pérez, (2007) quien señala que los instrumentos de evaluación pueden ser disímiles en relación con la información a recoger, el tipo de evaluación a aplicar, el tiempo, el contenido a evaluar y la materia, entre otros. Por lo anterior, es importante que las evaluaciones no se realicen con un solo instrumento ni de manera estandarizada.

Los resultados arrojados por este estudio de validez demuestran la importancia de que los estudiantes normalistas obtengan los conocimientos necesarios para el diseño e implementación de una diversidad de instrumentos de evaluación que les permita diversificar su proceso evaluativo y cumplir con el modelo de educación basado en competencias que establece un cambio en el rol que la educación superior tiene en la formación de profesionales que puedan enfrentar los nuevos desafíos y necesidades de la sociedad, junto con nuevas propuestas educativas que asignan un papel activo al estudiante en su proceso de aprendizaje, que apuntan a un tipo de aprendizaje profundo y significativo, y al logro de ciertas habilidades específicas y evaluables. Rosenbluth, Cruzat y Ugarte, (2015).

CONCLUSIONES

Es imprescindible contar con instrumentos evaluados y validados por expertos, validación que proporciona innumerables ventajas para evaluar y ajustar el instrumento de medición, en el caso que nos ocupa para el aprendizaje y enseñanza del diseño de instrumentos de evaluación en el proceso de formación de futuros docentes, para dar cabal cumplimiento a la serie de competencias genéricas y profesionales que el Plan de Estudios para la Formación de Maestros de Educación Primaria 2012 establece, pero en especial la competencia profesional: Emplea la evaluación para intervenir en los diferentes ámbitos y momentos de la tarea educativa (S.E.P., 2012: 11).

Por lo anterior, se concluye que el instrumento presentado es pertinente para evaluar y diagnosticar la metodología sobre el diseño de instrumentos de evaluación por estudiantes normalistas, tomando como base el juicio de expertos, el cual fue realizado por docentes con experiencia en el área y estudiantes normalistas, por ser estos últimos quienes se ven afectados en sus procesos de evaluación y que a la vez, serán los beneficiados por estos resultados al reconocer que “la importancia de unos instrumentos de evaluación adecuados radica en la trascendencia de las decisiones y las consecuencias que a partir de ellos se derivan, tanto personales como sociales” Muñiz y Fonseca, (2008), porque el conocimiento de diferentes instrumentos de evaluación permitirá diversificar el proceso evaluativo en el aula, buscando con ello, reconocer el nivel de conocimiento adquirido por los estudiantes con las actividades diseñadas en la planeación, aumentar las posibilidades de que el docente vaya mejorando en el análisis de su quehacer al reflexionar en lo que ha fallado y que le podrá permitir mejorar su desempeño.

También puede concluirse que el instrumento está redactado de manera adecuada y es comprensible por los potenciales usuarios, dado que el análisis de los jueces fue positiva en todos los ítems presentados. Diversos investigadores señalan la importancia de revisar muy

bien las preguntas y las opciones de respuesta para lograr alta validez de contenido que de acuerdo a Aliaga, (2007) consiste en determinar lo adecuado del muestreo de reactivos del universo de reactivos posibles; en este sentido, es una "medida" de lo adecuado del muestreo y, lograr la confiabilidad (o consistencia) precisión con que el test mide lo que mide, en una población determinada y en las condiciones normales de aplicación. Anastasi, (1982); Aiken, (1996), el coeficiente de confiabilidad es una correlación entre dos grupos de puntajes e indica el grado en que los individuos mantienen sus posiciones dentro de un grupo. Abarca valores desde 0 a 1. Cuanto más se acerque el coeficiente a 1, más confiable será la prueba Aliaga, (2007), tales como los resultados que se arrojaron en la V de Aiken descrita líneas arriba.

Visto lo anterior, es de señalar la importancia de la validez de pertinencia y redacción de este instrumento para la formación de los futuros docentes, a la vez es necesaria la realización de más investigaciones que lleven a la valoración de la importancia del aprendizaje, diseño, implementación de instrumentos de evaluación. Además, será de gran relevancia que existan más estudios validados por jueces expertos en la materia que permitan acrecentar los conocimientos de los estudiantes normalistas; así como, el desarrollo de habilidades profesionales relacionadas con el ejercicio de su actividad magisterial.

BIBLIOGRAFÍA

Aiken, L. (1996). Tests psicológicos de evaluación. México: Prentice-Hall.

Aliaga, J. (2007). "Psicometría: tests psicométricos, confiabilidad y validez. Recuperado de https://bit.ly/2AvTmNE

Anastasi, A. (1986). Los tests psicológicos. Madrid: Aguilar

Arribas, M. (2004). Diseño y validación de cuestionarios. Matronas Profesión 2004; vol. 5 (17): 23-29. Recuperado de http://ebevidencia.com/wp-content/uploads/2014/07/validacion_cuestionarios.pdf

Ato, M., López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. Anales de Psicología, 29(3), 1038-1059. https://dx.doi.org/10.6018/analesps.29.3.178511

Cabero, J. y Llorente, M. C. (2013), La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). En Eduweb. Revista de Tecnología de Información y Comunicación en Educación, 7 (2) pp.11-22. Disponible en http://tecnologiaedu.us.es/tecnoedu/images/stories/jca107.pdf

Calderón, C., & Gustems, J. (2013). Brown, S. y Pickford, R. (2013). Evaluación de habilidades y competencias en Educación Superior. Madrid: Narcea. . Anuario de Psicología, 43 (1), 129-130.

Camargo, I. y Pardo, C. (2008) Competencias docentes de profesores de pregrado: diseño y validación de un instrumento de evaluación. Bogotá, Colombia. Recuperado de http://www.redalyc.org/html/647/64770211/

Casanova, M. A. (2012), "El diseño curricular como factor de calidad educativa", Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, vol. 10, núm. 4, pp. 7-20. Recuperado de http://www.redalyc.org/pdf/551/55124841002.pdf

Castillo, S. y Cabrerizo, J. (2010) Evaluación educativa de aprendizajes y competencias. Madrid. Pearson

Cisneros, E. Jorquera, M. y Aguilar, Á. (2011) Validación de instrumentos de evaluación docents en el contexto de una Universidad Española. Voces y Silencios: Revista Latinoamericana de Educación, Vol. 3, No. 1, 41-55. Recuperado de http://Dialnet-ValidacionDeInstrumentosDeEvaluacionDocenteEnElCon-4054206.pdf

Cohen, R. & Swerdlik, M. (2001). Pruebas y evaluación psicológicas: Introducción a las pruebas y a la medición. (4ª ed.). México: Mc Graw Hill.

Consejo de Evaluación, Acreditación y Certificación de la Calidad de la Educación Superior Universitaria (2012). "Elaboración y validación de instrumentos de evaluación y competencias profesionales". Ministerio de Educación, Perú. Recuperado de https://www.sineace.gob.pe/wp-content/uploads/2017/03/03-Guia-elaboracion-Instrumentos-evaluacion.pdf

De Camilloni, A. (2000). La calidad de los programas de evaluación y de los instrumentos que los integran. La evaluación de los aprendizajes en el debate contemporáneo. 2ª reimpresión. México: Paidós, 67-92.

Díaz, M (2000). La formación de profesores en la educación superior colombiana: problemas, conceptos, políticas y estrategias. Bogotá: Instituto Colombiano para el Fomento de la Educación Superior

Dorantes, J., Hernández, J., y Tobón, S. (2016). Juicio de expertos para la validación de un instrumento de medición del síndrome de Burnout en la docencia. Ra Ximhai, 12 (6), 327-346.

Escobar, J. y Cuervo, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. En Avances en Medición, 6, pp. 27-36. Disponible en http://www.humanas.unal.edu.co/psicometria/files/7113/8574/5708/Articulo3_Juicio_de_expertos_27-36.pdf

Escurra, L. (1988). Cuantificación de la validez de contenido por criterio de jueces. Revista de Psicología. 6 (1-2), pp. 103-111.

Falchikov, N. (2005). Mejorar la evaluación a través de la participación de los estudiantes. Soluciones prácticas para ayudar Aprendizaje en educación superior y educación superior. Londres: RoutledgeFalmer.

Fernández, M. Cayssials, A. y Pérez, M. (2009). "Curso básico de psicometría: teoría clásica"

García Ñ., Fernández S.J. (2008). Procedimiento de aplicación de trabajo creativo en grupo de expertos. Energética. 29(2):46-50.

Hamodi, V. y López, A.. (2015). Medios, técnicas e instrumentos de evaluación formativa y compartida del aprendizaje en educación superior. Perfiles educativos 2015; 37:146-46. DOI: 10.1016/j.pe.2015.10.004. Recuperado de http://www.elsevier.es/es-revista-perfiles-educativos-85-articulo-medios-tecnicas-e-instrumentos-evaluacion-S0185269815000100#bib0270

Hernández - Mosqueda, J., & Tobón - Tobón, S., & Guerrero - Rosas, G. (2016). Hacia una evaluación integral del desempeño: las rúbricas socioformativas. Ra Ximhai,12 (6), 359-376.

Hernández, J. S., Tobón, S., González, L., & Guzmán, C. (2015). Evaluación socioformativa y rendimiento académico en un programa de posgrado en línea. Paradigma, 36(1), 30-41. Recuperado de https://goo.gl/gWcYTx

Hernández, R. y otros (1999) "Metodología de la investigación".McGRAW-HILL

Hyrkäs, K., Appelqvist-Schmidlechner, K & Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. International Journal of nursing studies, 40 (6), 619 -625.

Ibarra, M.S. y Rodríguez, G. (2011). Los procedimientos de evaluación. En Rodríguez e Ibarra e-evaluación orientada al e-aprendizaje estratégico en educación superior. Madrid: Narcea.

McGartland, D. Berg, M., Tebb, S. S., Lee, E. S. & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. Social Work Research, 27 (2), 94-104

Montero, I. y León, O. (2002) "Clasificación y descripción de las metodologías de la investigación en Psicología". Madrid, España, Recuperado de http://www.aepc.es/ijchp/articulos_pdf/ijchp-53.pdf

Muñiz, J. y Fonseca, E. (2008) Construcción de instrumentos de medida para la evaluación universitaria. Recuperado de: http://reined.webs.uvigo.es/ojs/index.php/reined/article/viewFile/40/35

Penfield, R.D. y Giacobbi, P.R., Jr. (2004). Applying a score confidence interval to Aiken's item contentrelevance index. Measurement in Physical Education and Exercise Science, 8(4), 213-225.

Rhodes, G. y Tallantyre, F. (2013) Evaluación de las habilidades básicas. Evaluar en la universidad. Problemas y enfoques pp. 129-143

Rosenbluth, A; Cruzat-Mandich, C. y Ugarte, M. L. (2015). Metodología para validar un instrumento de evaluación por competencias en estudiantes de psicología. Univ. Psychol. [online]. 2016, vol.15, n.1, pp.303-314. ISSN 1657-9267. http://dx.doi.org/10.11144/Javeriana.upsy15-1.ppmp.

Ruiz, J. (2005). La evaluación de la docencia en los planes de mejora de la universidad. Educación, XXI (8), 87-102

Sáez, JM. (2017)."Investigación educativa. Fundamentos teóricos, procesos y elementos prácticos" (cap. 6

Sanmartí, N. (2010) "10 ideas claves. Evaluar para aprender". Madrid, España. Editorial GRAÓ.

Secretaría de Educación Pública (2012) Acuerdo por el que se establece el Plan de Estudios para l Formación de Maestros de Educación Primaria. Ciudad de México, México.

Sladogna, G. M. (2000). La recentralización del diseño curricular. El perfil profesional y la definición de competencias profesionales. Proyecto INET-GTZ. Cooperación Alemana para el Desarrollo. Buenos Aires: Instituto Nacional de Educación Tecnológica.

Tobón, S. (2013) Evaluación de las competencias en la educación básica (2°. Ed). México: Santillana

Tobón, S. (2013b). Formación integral y competencias. Pensamiento complejo, currículo, didáctica y evaluación (4a. Ed.). Bogotá: ECOE

Tobón, S. (2017, octubre). Metodología de la evaluación socioformativa. En L. G. Juárez-Hernández (Moderador), II Congreso Internacional de Evaluación del Desempeño, Valora. Congreso conducido por el Centro Universitario CIFE, Cuernavaca, México.

Zabala, A., & Arnau, L. (2007). Cómo aprender y enseñar competencias. Barcelona: Graó. Recuperado de http://goo.gl/bXkiYg