

PONENCIA: DETERMINACIÓN DEL ÍNDICE DE CONSISTENCIA INTERNA DE LA PRUEBA PEDAGÓGICA.

ELABORÓ: M.C. Sergio Trejo Mancillas.

Santa María del Oro, El Oro, Durango.

21 de enero de 2014

Resumen

El presente trabajo de investigación contiene el proceso a seguir en la validación de la prueba pedagógica mediante el índice de consistencia interna el cual es deseable que se obtenga a 0.50. El contar con instrumentos confiables, permite a la metodología de la investigación utilizar mecanismos que hagan posible medir el impacto de la variable independiente sobre la variable dependiente en la investigación experimental cuantitativa.

Durante el proceso de piloteo se utilizó una muestra intencionada con 20 alumnos del VI semestre de la Licenciatura en Educación Preescolar de la Escuela Normal Profesor Carlos A. Carrillo de Santa María del Oro. El Oro, Durango en el periodo de Agosto de 2012 a Junio de 2013.

La prueba pedagógica validada se conformó por 21 ítems, los cuales comprendían un índice de dificultad del reactivo desde 0.20 a 0.80 para obtener un índice promedio de dificultad de la prueba de 0.51, asimismo en el diseño del ítem se atendieron orientaciones propuestas por el Ceneval y por la Dirección General de Bachillerato del estado de Veracruz, de esta forma los reactivos se ubican en el primer nivel cognoscitivo (Conocimiento) conforme la taxonomía de Benjamín Bloom.

Palabras Claves: Validez, confiabilidad e índice de dificultad del reactivo (P)

DELIMITACIÓN DEL TEMA DE ESTUDIO

Uno de los principales problemas que se tienen en el proceso de toda investigación es contar con la validez de instrumentos, los cuales permiten tener cierto grado de certeza en su confiabilidad y validez en su aplicación para la medición de variables.

Así un instrumento de medida es válido si mide lo que supone medir (validez de constructo), si representa lo que dice representar (validez de contenido), y si sirve para lo que dice servir (validez predictiva).

Es importante respetar los fines del instrumento por lo que para ello se habrá de exigir un tipo de validez, así la validez de contenido es indispensable para el diagnóstico y el tratamiento psicológico.

La Validez de constructo es imprescindible para los trabajos de investigación científica y la validez predictiva es imprescindible para trabajo de predicción pronóstico o selección de personas.

Generalmente la validez se calcula correlacionando las puntuaciones del instrumento con un criterio externo. Los criterios externos pueden ser:

- a) Otro instrumento con validez ya conocido.
- b) El juicio de expertos.
- c) Algún comportamiento que guarda alta relación con el constructo que dice medir.
- d) Cálculo del índice de dificultad de la prueba pedagógica.

De los cuatro métodos descritos con anterioridad únicamente nos referiremos al índice de dificultad del reactivo en la prueba pedagógica.

El determinar la validez de constructo o la de contenido permite al investigador contar con instrumento confiable para poder someter a prueba de hipótesis la variable independiente, sin embargo existen ciertas contradicciones sobre el hecho de que un instrumento de medición como lo es la prueba pedagógica puede tener confiabilidad y validez, pues autores como Briones (2001, p.155) consideran que una de las desventajas que tienen el método de la repetición de la prueba pedagógica es el hecho de que el alumno por aprendizaje previo pueda modificar o alternar la confiabilidad de la misma, pues al aplicar la misma prueba pedagógica al estudiante en dos momentos diferente, existe el riesgo de que el alumno recuerde la información de las respuestas y preguntas durante la primer aplicación del

instrumento; esta situación nos orilla a realizar el siguiente planteamiento del problema:

- **¿Cuál es el procedimiento a desarrollar en la obtención del índice de dificultad del reactivo en la prueba pedagógica?**
- ¿Qué tipo de nivel cognoscitivo debe tener cada uno de los ítems en el diseño del instrumento?
- ¿Qué dificultades se presentan durante el proceso del piloteo de la prueba pedagógica?
- ¿Cuál es el índice de consistencia interna que debe tener la prueba pedagógica?

Objetivo general

- Determinar el procedimiento a desarrollar en la obtención del índice de dificultad del reactivo en la prueba pedagógica.

Objetivos específicos

- Diseñar el instrumento de medición, determinando el nivel cognoscitivo que debe tener cada uno de los ítems.

- Conocer las dificultades que se presentan durante el proceso de piloteo de la prueba pedagógica.
- Medir el índice de consistencia interna de la prueba pedagógica.

Justificación

El poder determinar el índice de consistencia interna que debe tener una prueba pedagógica contribuirá a la metodología de la investigación contar con instrumentos que permitan medir el impacto que tiene una variable independiente, sobre la variable dependiente en metodología de la investigación.

Revisión de literatura

En México la mayoría de las pruebas pedagógicas que se aplican están estandarizadas por el *Centro Nacional de Evaluación (Ceneval)* quien es responsable de medir los niveles de aprovechamiento en educación superior.

-El contenido: este debe corresponder al perfil referencial las explicaciones del reactivo de la prueba, de tal manera que el ítem pueda proporcionar información acerca del conocimiento y habilidades que tiene el estudiante sobre la prueba.

-La base: incluir solamente una idea al elaborar el reactivo para que el estudiante pueda demostrar el nivel de dominio que tiene.

-Las opciones de respuesta: verificar la correspondencia gramatical entre la base y el inicio de cada una de las opciones, cuidar que guarden equilibrio entre los elementos gramaticales como puede ser el género, número, tiempo verbal, sintaxis, etcétera.

-Sobre los distractores: incluir en los distractores los errores más comunes de los alumnos en lugar de enunciados descartables por sentido común o inferencia lógica (Centro Nacional de Evaluación,[Ceneval], s.f. pp. 5-6).

Es sumamente importante que para poder lograr la validez de las pruebas pedagógicas partir de algunas concepciones como:

- Validez: Cuando se habla de validez de un instrumento se quiere significar la propiedad por la cual un instrumento mide realmente aquello que pretende medir, o por tanto, es válido porque conduce a conclusiones válidas (Martínez,[s. f.], Validez, párr.1).
- Confiabilidad: es la propiedad según la cual un instrumento aplicado a los mismos fenómenos, bajo los mismos condiciones, arroja resultados congruentes (Martínez, [s.f.], Confiabilidad, párr. 1).
- Consistencia interna: propiedad del instrumento según el cual, cada parte del mismo mide de la misma manera que las otras partes (Martínez, [s.f.], consistencia interna, párr. 1).

En los grupos pequeños se puede tomar el 50%, es decir, la mitad del grupo pertenecerán al grupo superior y la otra mitad del grupo pertenecerá al grupo inferior por lo que el índice estadístico se calcula mediante la siguiente fórmula:

$$P = \frac{Up + Lp}{U + L}$$

$$D = \frac{Up - Lp}{U}$$

Generalmente el valor de P es conocido como índice de dificultad del reactivo, y el valor de D se le conoce como el índice de discriminación.

Usualmente el índice de dificultad del reactivo tiene un rango de 0.00 a 1.00. Un reactivo generalmente requieren de un índice de 0.50 es decir el reactivo no debe ser ni muy fácil de contestar ni al mismo tiempo muy difícil.

Sin embargo generalmente en las pruebas pedagógicas se incluyen reactivos con un rango bastante estrecho, sea éste de 0.20 y aunque casi siempre se incluyen ítems en la prueba bastante difíciles y otros muy fáciles de resolver (Capítulo cuatro. Análisis de reactivos y estandarización de pruebas, [s. f.], p. 66).

Por su parte el índice de discriminación del reactivo de una prueba pedagógica (D), permite conocer entre quienes obtienen altas y bajas calificaciones en una prueba, por lo general se considera aceptable un valor de “D” igual a 0.30 de tal forma que incluso en ocasiones se pueden aceptar un valor de D aún un poco más inferior a 0.30.

La siguiente tabla expresa los índices medios óptimos de la dificultad de los reactivos para prueba con reactivo de opción múltiple.

Índice Medio Óptimo de	
Números de opciones	Dificultad (p.)
2	0.83
3	0.77
4	0.74
5	0.69
Abierta	0.50
(Ensayo, respuesta breve)	

Fuente: Elaborado con datos proporcionados por F. M. Lord, Psychometrika, 17, (1952), pp. 181- 194. (Como se citó en capítulo cuatro, [s. f.]. Análisis de reactivos y estandarización de pruebas).

METODOLOGÍA

Durante el proceso de validación del instrumento se habrá de diseñar el instrumento, además de seleccionar los participante para el piloteo, y determinar el índice de consistencia interna de la prueba.

Elección y argumentación del método

Por el enfoque cuantitativo que tiene la presente investigación de la validación de la prueba pedagógica, podemos encontrar que está basada en la inducción

probabilística del positivismo lógico, la medición es penetrante y controlada e incluso es objetiva de tal forma que la validación de la prueba nos puede inferir que efectivamente cuales reactivos son los que deben tomarse en cuenta para su aplicación (Herrera, (s. f.) p. 6).

MUESTRA: EL muestreo será intencionado por lo que el investigador habrá de seleccionar al grupo del VI semestre de la Licenciatura en Educación Preescolar para efectuar el piloteo de la prueba.

En este tipo de muestras: El investigador selecciona los elementos que a su juicio son representativos, lo cual exige al investigador un conocimiento previo de la población que se investiga para poder determinar cuáles son las categorías o elementos que se pueden considerar como tipo representativo del fenómeno que se estudia (Tamayo, 2004, p.78).

Elección de participantes en la investigación

En este caso se seleccionarán 20 alumnos para el piloteo de la prueba de La Educación en el Desarrollo Histórico de México II por lo que deberán de cubrir los siguientes requisitos:

- a). Ser alumno de la Escuela Normal Profesor Carlos A. Carrillo.
- b) Sexo masculino o femenino con edades comprendidas de 18- 22 años
- C). Tener máximo un año de haber cursado la asignatura anteriormente mencionada.

Asimismo se seleccionarán tres docentes quienes deberán de tener la experiencia en la impartición al menos en dos ocasiones de esta asignatura.

Elección, descripción y argumentación de técnicas de obtención empírica.

Una vez piloteado el instrumento, se aplicará al IV semestre de la Licenciatura en educación preescolar para determinar el índice de dificultad del reactivo (P) y el índice de discriminación de la prueba pedagógica (D) para valorar la consistencia interna a través de la fórmula:

$$P = \frac{Up + Lp}{U + L} \qquad D = \frac{Up - Lp}{U}$$

Donde: Up y Lp representan la cantidad de sujetos tanto en el grupo superior como el grupo inferior que conteste bien el reactivo.

P = Índice de dificultad del reactivo.

D = Índice de discriminación del reactivo.

La U= representa la cantidad de sujetos que hay en el grupo superior la cual debe ser igual a L que es la cantidad de sujetos que existen en el grupo inferior.

Diseño del instrumento

Por tratarse de una ponencia y situaciones de espacio solamente se presenta un extracto del diseño de la prueba pedagógica.

Después de realizar una revisión al material de elaboración de reactivos de ceneval, [s, f.] se encontró que estos tienen seis niveles: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación (p. 11).

Para los fines de la presente investigación se optó por que estos fueran elaborados para medir “conocimiento” en consecuencia el diseño del reactivo debería de tener verbos como: definir, identificar, listar, reproduzca, pronuncie, declare, cite, etiquete, relacione, nombre, etc.

De igual forma se deben responder a cuestionamientos como: ¿Qué enunciado define el término?, ¿En qué año se descubrió?, ¿Cuál es la forma?, ¿Cuál es el objeto de estudio de...? (Bloom,[s.f.]. Como se citó en ceneval, [s. f.] p. 16).

Durante la elaboración de los ítems se prestó atención a los aspectos relacionados con: *el contenido del reactivo* que hiciera posible la verificación del conocimiento del alumno. De igual manera en el presente examen se utilizaron 11 reactivos de respuesta breve, evitando ofrecer opciones al alumno para poder comprobar su conocimiento. (Dirección General de Bachillerato [DGB]. Manual Para la Elaboración de Reactivos, [s. f.], p. 24). Tal como se muestra enseguida:

1. ¿Qué tipo de participación tuvieron los maestros en México durante el periodo revolucionario?..... ()

Asimismo se prestó especial atención a los 5 reactivos que se incluyeron en la segunda parte del examen la cual contemplaba preguntas de correspondencia.

II. Relacione ambas columnas.

()	Luz María Serradell logró que las educadoras formaran parte del SNTE.	1. 1934,1940 período de Lázaro Cárdenas. 2. 1935.
()	La reforma del artículo 3° constitucional establece la educación socialista en el periodo....	3. 1938.
()	Educadora cuya conferencia en 1937 declaraba que...“imprimir a la educación en el jardín de niños el sello de nacionalismo, de vitalidad, de utilidad y servicio social”.	4. Rosaura Zapata. 5. Estefanía Castañeda.
()	General que asume la presidencia de la República en 1940.	6. Ávila Camacho.

		7. Lázaro Cárdenas.
()	Durante este periodo el programa de los jardines estableció diferencias para tres grados de preescolar.	8. 1940-1946.

En los 5 reactivos de Falso o Verdadero, sólo es posible que el alumno pueda elegir una de estas respuestas (DGB, Manual para la elaboración de reactivos, [s. f.], p. 26).

El 13 de diciembre de 1916 no fue celebrado el debate de la 12va. Sesión ordinaria del artículo 3° constitucional....._____

Análisis y discusión de resultados

La prueba pedagógica quedó conformada por 21 ítems con un índice promedio de dificultad de la prueba de 0.51.

Las principales dificultades durante el piloteo de la prueba fue el hecho de no contar con grupos paralelos así como no tener maestros suficientes en la especialidad del examen para su validación.

Para determinar el índice de dificultad de cada ítem se utilizó la fórmula descrita anteriormente y los reactivos de cada alumno, obsérvese la tabla No. 1

Examen: La Educación en el Desarrollo Histórico de México II

Tabla No. 1

Pre Test: 12-12-12
II Semestre: Licenciatura en Educación Preescolar

No. Progresivo.	Ítems																					No. De reactivos	Califica- ción
	1	5	6	8	11	12	13	16	17	18	20	21	22	23	24	25	26	27	28	29	30		
17	0	0.5	1	1	1	0	0	0.5	1	0.5	1	1	1	1	1	1	0	1	1	1	1	18.5	32
4	0.5	1	0	1	0	0	1	0.5	1	1	1	0	1	1	1	1	1	0	1	1	1	18	31
14	0	0	1	1	1	1	1	0.5	0.5	1	1	0	0	1	1	0	1	0	1	1	1	17.5	30
16	1	0.5	1	1	1	0.5	1	0.5	1	1	1	1	1	0	0	0	1	0	1	1	1	17.5	30
2	0	0.5	0.5	0	1	0	1	0.5	0.5	1	1	0	1	1	1	1	1	0	1	1	1	16	28
18	0	0.5	1	0	1	0	1	0.5	1	1	1	0	1	0	1	1	1	0	1	1	1	16	28
8	0.5	0	1	0	1	0	0.5	0.5	0	1	1	1	0	1	0	0	1	1	1	1	1	15.5	27
6	1	0.5	0	0	1	1	1	0.5	0	0.5	0	0	1	0	1	1	1	0	1	1	1	15	25
15	1	0.5	1	0	0	1	1	0.5	0.5	0.5	1	0	1	0	1	1	0	0	0	1	1	14	24
19	0	0	0.5	0	1	0	0.5	0.5	0.5	0.5	1	0	1	0	1	1	1	0	1	1	1	13.5	23
21	0	0	0	0	0	0	0	0	0.5	1	1	1	1	1	1	1	1	0	1	1	1	13.5	23
10	0	0	1	1	1	0	0.5	0.5	1	0.5	1	0	0	1	0	0	0	1	1	1	0	12.5	22
11	0	0	0	0	1	0	1	0.5	0.5	1	1	0	1	0	1	1	1	0	1	0	0	12	21
13	0	0.5	0	0	0	1	1	0	1	0.5	0	0	0	1	1	1	0	0	1	1	1	12	21
9	0	0.5	1	0	0	0	1	0.5	0.5	0	0	0	0	0	0	0	1	0	1	1	1	10.5	18
3	0	0	0	0	1	0	1	0	0	0.5	1	0	0	1	1	0	0	0	1	0	1	9.5	17
20	0	0	0	0	1	0	0.5	0	0	0.5	0	0	0	1	1	1	0	1	1	0	0	8	14
1	0	0	0	0	0	0	0	0	0	0.5	1	0	0	1	1	0	1	0	0	0	1	7.5	13
12	0	0	0	0	1	0	0.5	0	0	0.5	1	0	0	0	0	1	1	0	0	0	0	7	12
7	0	0	0	0	1	0	0	0	0	0.5	1	0	0	1	1	0	0	0	0	0	0	6.5	11
5																							
Índice de dificultad del ítem	0.20	0.25	0.45	0.25	0.70	0.22	0.67	0.32	0.47	0.67	0.80	0.20	0.50	0.60	0.75	0.60	0.65	0.20	0.80	0.70	0.75	0.51	
Índice de discriminación del reactivo	0.40	0.30	0.50	0.30	0.20	0.25	0.25	0.35	0.25	0.27	0.20	0.20	0.60	0.20	0.10	0.20	0.30	0.00	0.20	0.60	0.50	0.29	

Códigos:

1= Respuesta Correcta

0= Respuesta incorrecta

0.5= A criterio del maestro, sólo el 50% de la respuesta es correcta.

Nota: Faltan en la tabla, 9 ítems con su respectivo valor de P, los cuales se descartaron por tener valores por debajo de 0.20 y por encima de 0.80, se obtuvo un valor de $P=0.51$ y el alumno número 5 de la lista no presentó examen, razón por la cual aparece en blanco este espacio.

BIBLIOGRAFÍA

Briones, G. (2001). *Métodos y técnicas de investigación para las Ciencias Sociales*. México, D.F.: Editorial trillas.

Ceneval. (s.f.). *Ejemplos de reactivos por nivel cognoscitivo*, pp. 1-15.

Dirección General de Bachillerato de la Secretaría de Educación de Veracruz. (s. f.). Manual para la elaboración de reactivos. Bajo el enfoque por competencias. Recuperado del sitio de internet de:

[http://art3resp.zapto.org:8080/Maestros/Maestros/Manual%20PARA%20LA%20ELABORACION%20DE%20REACTIVOS%20\(DggGB\).pdf](http://art3resp.zapto.org:8080/Maestros/Maestros/Manual%20PARA%20LA%20ELABORACION%20DE%20REACTIVOS%20(DggGB).pdf).

Herrera, J. (s.f.). *Investigación Cuantitativa*. Recuperado de: <http://JuanHerrera.files.wordpress.com/2008/11/investigación-cuantitativa.pdf>.

Martínez, M. A. (s. f.). *Validación (Validez y confiabilidad)*. Recuperado de: http://www.arnaldomartínez.net/enfermeria/validez_y_confiabilidad.pdf.

Tamayo, y T.M.(2004). *El proceso de la investigación científica*. México, D.F.: Editorial Limusa.