



ISSN: 2448-6574

Diseño de una prueba objetiva para la evaluación de aprendizajes adquiridos de la materia de Psicología Evolutiva de la Infancia

Cuauhtémoc Gerardo Pérez López

cgperez@upn.mx

María Yolanda Quiroz Arce

maquiroz@upn.mx

Aristarco Noel Méndez Lechuga

arimex2@hotmail.com

Alba Yanalte Álvarez Mejía

yalvarez@upn.mx

María Imelda González Mecalco

piowacwe@yahoo.com.mx

Resumen

El presente documento informa acerca del diseño y aplicación piloto de una prueba de evaluación objetiva de los aprendizajes de contenidos revisados en el curso Psicología evolutiva de la infancia que se imparte en la Licenciatura en psicología educativa en la Universidad Pedagógica Nacional. Para tal fin se diseñaron 30 reactivos con tres opciones –la respuesta correcta y dos distractores- que evalúan los aprendizajes esperados de la asignatura mencionada. El instrumento se aplicó a 332 estudiantes quienes tomaron ese curso en el semestre previo. Se hizo el cálculo de los indicadores psicométricos de los reactivos: el índice de discriminación y la facilidad. Además se obtuvo el porcentaje de respuestas en los distractores. Los hallazgos se discuten a la luz de lo que aportan al proceso de evaluación construcción del instrumento de este y el resto de los cursos –siete en total- que conforman la línea curricular Desarrollo y aprendizaje. Asimismo, los resultados serán la base para dar inicio a un posterior trabajo de seguimiento de los aprendizajes adquiridos por los estudiantes durante



ISSN: 2448-6574

su formación, en el que la evaluación objetiva sea solo uno más de los procedimientos que permitan alcanzar el estudio de los procesos con una aproximación global de la evaluación.

Palabras clave: Evaluación educativa, evaluación de los aprendizajes, pruebas objetivas para evaluar los aprendizajes

Introducción

Los cursos de la Licenciatura en Psicología Educativa (LPE) de la Universidad Pedagógica Nacional se organizan en tres Fases y siete Líneas Curriculares (LC). Una de las líneas se denomina Desarrollo y Aprendizaje, su propósito formativo es la comprensión y análisis de los procesos de desarrollo (cognitivos, afectivo, relacionales) y de aprendizaje, en relación con los cambios derivados de la participación en los contextos educativos.

La LC está constituida por siete asignaturas, tres de ellas se centran en el estudio del desarrollo evolutivo de los individuos. En el primero se estudia la infancia, en el segundo la adolescencia y, en el tercero, la adultez y la vejez. El primero de ellos, y sobre el cual versa este trabajo, tiene como propósito caracterizar los procesos de cambio en el desarrollo humano (físico, psicológico y social), comparando las teorías, conceptos básicos y diseños de investigación en las áreas del desarrollo cognitivo, afectivo y de lenguaje. Además de discutir y reflexionar sobre las implicaciones educativas de las teorías revisadas.

El presente es parte de un trabajo de investigación más amplio, que incluye analizar los aprendizajes de los cursos de Psicología evolutiva de la adolescencia y Psicología evolutiva de la adultez y vejez. Este documento presenta el proceso seguido en el diseño y aplicación de una prueba objetiva de evaluación de los aprendizajes del curso Psicología Evolutiva de la Infancia. En esta primera etapa se tomó una muestra de estudiantes de segundo semestre de la LPE, quienes cursaron dicha materia en primer semestre. El trabajo tiene como propósito reportar el comportamiento del instrumento para evaluar los aprendizajes de los estudiantes sobre contenidos de desarrollo evolutivo de la infancia, incluidos en el programa de estudio.



ISSN: 2448-6574

Revisión teórica

De manera regular, la evaluación en la educación superior suele centralizarse en instancias ajenas o lejanas a los docentes y las actividades cotidianas de enseñanza. Por lo cual las conclusiones y decisiones derivadas de dicha evaluación se enfocan en los resultados de los estudiantes, tales como la reprobación o la eficiencia terminal, dejando de lado los procesos formativos.

Desde la concepción de rediseño del plan de estudios de la Licenciatura en Psicología Educativa, se plantea, entre otras acciones, la necesidad de dar seguimiento a los procesos de aprendizaje de los estudiantes. Este procedimiento permitirá dar validez y confiabilidad a los nuevos programas de estudio (Hernández, Pérez, Martínez, Bollás y Dzib, 2009).

Hacer el seguimiento y la evaluación de los planes de estudio de las instituciones de educación superior, nacionales y extranjeras, se reconoce como un elemento esencial de mejora y ajuste de los procesos educativos a fin de conseguir las metas u objetivos previstos. Es por ello que se hace necesario el desarrollo e incorporación de una metodología de evaluación curricular que incluya elementos suficientes para comprender la situación del programa a evaluar. En el presente trabajo se lleva a cabo una propuesta la cual se integra por elementos de evaluación interna centrada en los aprendizajes escolares formativos de la carrera en Psicología Educativa de la Universidad Pedagógica Nacional (Licenciatura en psicología educativa, 2011).

La evaluación educativa puede ser vista desde distintas perspectivas teóricas. Así, Scriven (1967) formuló las funciones de la evaluación como sumativa y formativa. La primera con un carácter más bien de control y certificación de los resultados, en tanto que la formativa como medio para mejorar el objeto de evaluación.

Mateo (2005) afirma que existen dos aproximaciones evaluativas: la evaluación centrada en los resultados y aquella orientada al estudio de los procesos, en la primera se hace uso de recopilación de información que permite análisis fragmentarios y parciales, la segunda permite una aproximación global a la información.



ISSN: 2448-6574

Para llevar a cabo la evaluación de los aprendizajes del curso Psicología evolutiva de la infancia, se revisaron diversos procedimientos para la recogida de información. Lukas y Santiago (2009) presentan una clasificación que da cuenta de técnicas de corte cuantitativo y otras que se dirigen a una obtención de información más bien cualitativa, si bien se inicia con una prueba objetiva que informe en su primera etapa del aprendizaje de contenidos de la materia de Psicología Evolutiva de la Infancia.

Con este trabajo se pretende contar con procesos de evaluación y seguimiento que permitan establecer rutas de mejora que impacten no solo en la adecuación y actualización de los programas y contenidos específicos de la materia, sino también que permitan instrumentar propuestas didácticas para innovar la práctica docente que posibilite al profesor proponer estrategias para generar la adquisición de aprendizajes significativos.

En el diseño de un sistema de seguimiento y desarrollo curricular de la licenciatura en psicología educativa de la UPN se consideran distintos niveles, objetos de estudio y participantes (Hernández, Pérez, Martínez, Bollás y Dzib, 2009). Interesa particularmente atender los procesos de programación de los cursos (descripción de los procesos de participación en el desarrollo de las líneas curriculares, elaboración de los programas y unidades de trabajo, manejo e impartición del programa y seguimiento de las estrategias docentes y de evaluación de los aprendizajes, así como revisión y actualización de los programas de estudio); prácticas docentes (caracterizar las distintas formas de ejercicio docente e identificar los rasgos formativos que tiene la docencia en los estudiantes); y analizar los indicadores de desempeño institucional (trayectorias escolares, retención, reprobación y abandono escolar, eficiencia terminal y titulación, entre otros).

En este marco, el objetivo del presente es dar cuenta del diseño y aplicación de una prueba objetiva para evaluar el aprendizaje de contenidos del curso Psicología evolutiva de la infancia, en estudiantes de la licenciatura en psicología educativa de la Universidad Pedagógica Nacional, Unidad Ajusco.

MÉTODO

Dado el interés por informar de la construcción de la prueba y los resultados de su primera aplicación (estudio piloto), el presente es un estudio descriptivo exploratorio.

Participantes

Participaron 332 estudiantes inscritos en el segundo semestre de la Licenciatura en psicología educativa. Con un rango de edad entre los 19 y 35 años. En el momento de la aplicación cursaban la asignatura Psicología evolutiva de la adolescencia.

Instrumento

Se utilizó una prueba objetiva de evaluación de contenidos curriculares, construida *ex profeso*. Se compone de 30 reactivos que evalúan los contenidos de las cinco unidades que integran el curso. La distribución de reactivos se deriva en función del contenido de cada unidad, con la finalidad de no construir reactivos repetidos o carentes de sentido por necesitar un contexto más amplio que los identifique (ver tabla 1).

Tabla 1. Distribución de reactivos por unidad del curso

	Nombre de la unidad	Reactivos
1	La psicología evolutiva: objeto de estudio, modelos explicativos	12
2	El desarrollo cognitivo del niño	6
3	Desarrollo afectivo	4
4	Desarrollo del lenguaje y habilidades comunicativas	6
5	Perspectiva sociocultural del desarrollo	3

Comentario especial requiere la distribución desigual de reactivos en las distintas unidades del curso. Ello se debe al peso específico que para los docentes asignan a los contenidos que se revisan al inicio del curso; estos, a decir de los profesores, son temas que serán medulares en la comprensión de los conceptos posteriores.

Para el diseño de los reactivos se decidió utilizar el procedimiento que se ha seguido en la licenciatura para elaborar los reactivos que conforman el examen general de conocimientos del plan 09 (Pérez, Bollás, González, Álvarez y Castro, 2014). De esa manera, se tomaron en cuenta los criterios que proponen Haldayna, Downing y Rodríguez (2002); Moreno, Martínez y



ISSN: 2448-6574

Muñiz (2004) y CENEVAL (2000; 2011). Asimismo se optó por el diseño de reactivos con tres opciones de respuesta. Lo anterior fue decidido a partir de los informes de Baghaei y Amrahi (2011), Bruno y Dirkzwager (1995), Haladyna y Downing (1993) y Rodríguez (2005), en los que se menciona que en los exámenes de elección múltiple, tres opciones distractoras son suficientes.

Por último conviene resaltar que para el diseño de este instrumento se supone el modelo secuencial. De este modo, los reactivos que se usaron en esta fase del estudio serán el insumo principal de las siguientes versiones del instrumento en futuras aplicaciones.

Procedimiento

La aplicación se realizó de manera colectiva en los salones de la universidad, en el horario de clases ordinarias, en un tiempo promedio de 30 minutos. Se concertó con el profesor el día y la hora de aplicación. Es pertinente mencionar que los estudiantes cursaron esa materia durante el semestre anterior. Es decir, había transcurrido un par de meses desde que se concluyó el curso. La actividad se comenzó quince minutos después de iniciada la sesión, esto con el fin de garantizar que la mayoría de los estudiantes estuviera ya en el aula.

RESULTADOS

Las respuestas de los participantes se registraron en una base de datos para poder analizarlas en dos fases: a) calificar la respuesta como correcta o incorrecta; con esta información se calcularon los indicadores psicométricos de la prueba y, b) determinar la frecuencia de respuesta para cada opción del reactivo. Con estos datos se pudo identificar qué reactivo requiere ser revisado, adecuado o sustituido.

En la tabla 2 se observa la distribución de los reactivos de acuerdo con su índice de discriminación. Es importante mencionar la manera muy homogénea en la que se distribuyen los reactivos en los rangos establecidos en la literatura especializada (Backhoff, Larrazolo y Rosas, 2000). Se tiene casi en partes iguales reactivos buenos, reactivos regulares y reactivos pobres. En consecuencia, es posible conservar todos para las siguientes fases del estudio. Empero es indispensable revisar de manera general once de ellos y hacer una revisión a

Debates en Evaluación y Currículum/ Congreso Internacional de Educación Evaluación 2016 / Año 2, No. 2, Septiembre de 2016 a Agosto de 2017/

profundidad en los otros diez. Es conveniente mencionar que el índice de discriminación general de la prueba fue de 0.24. Este valor, no obstante ubicarse en el rango de regular, debe considerarse como bueno dado que es la primera ocasión que se aplica la prueba.

Tabla 2. Poder de discriminación de los reactivos según su valor D

Valor D	Frecuencia	Calidad del reactivo
> 0.39	1	Excelente
0.30 – 0.39	8	Buena
0.20 – 0.29	11	Regular
0.0 - 0.19	10	Pobre
> -0	0	Pésima

En relación con la facilidad (ver tabla 3), se observa que tres de cada cuatro reactivos se ubican en los rangos ideales de este indicador. La mayor cantidad de reactivos que tienden a difíciles y uno muy difícil hace que la facilidad general de la prueba sea 0.45. Este no es el puntaje óptimo (0.50), empero sí es un puntaje cercano.

Tabla 3. Valor de facilidad de los reactivos

Facilidad	Frecuencia	Calidad del reactivo
91 – 100	0	Muy fácil
0.73 – 90	2	Tiende a fácil
0.72 – 0.28	23	Buen reactivo
0.27 - 0.11	4	Tiende a difícil
0.10 - 0.0	1	Muy Difícil

Con los puntajes de ambos indicadores es posible decir que el comportamiento general del instrumento fue bueno; como se dijo antes debe trabajarse algunos reactivos, tal vez otros debían ser reformulados totalmente, pero la mayor parte de ellos, como se dijo ya, podrá conservarse (ver tabla 4).

Tabla . Distribución de reactivos que deberán revisarse por el criterio de facilidad

Discriminación	Facilidad	Frecuencia
0.01 - 0.10	< 0.28	0
0.01 - 0.10	> 0.73	0
0.11 - 0.20	< 0.28	1
0.11 - 0.20	> 0.73	1
0.21 - 0.30	< 0.28	1
0.21 - 0.30	> 0.73	0

0.31 - 0.40	< 0.28	1
0.31 - 0.40	> 0.73	0

Para completar la información técnica del reactivo, se requiere conocer la distribución porcentual de respuestas en los distractores. La distribución ideal será que cada uno de los distractores tengan la mitad de las respuestas incorrectas. En la tabla 4 se muestra la distribución de respuestas por reactivo y opción, en tanto en la tabla 5 se observa el tipo de distribución en el total de reactivos. Se consideró distribución homogénea cuando la diferencia fue menor a 5 puntos porcentuales. Se observa que hay casi el doble de reactivos en los que alguna de las opciones tiene una frecuencia de respuestas mayor; no obstante eso supone un 36% de reactivos cuyas opciones son igualmente atractivas para los sustentantes. Debe mencionarse que en cuatro de los cinco que son difíciles, dicha distribución fue homogénea. La tercera columna registra la frecuencia de cuántos reactivos tuvieron un porcentaje menor de 10% en alguna de las opciones. Por supuesto, esto sucedió en los dos reactivos más fáciles, además de tener una distribución heterogénea.

Tabla 4. Distribución de frecuencia y porcentaje por reactivo y opción

Reactivo	Opción A		Opción B		Opción C		No respuesta	
	Frec	%	Frec	%	Frec	%	Frec	%
1	11	3.3	43	13	277	83.4	1	0.3
2	148	44.6	151	45.5	27	8.1	6	1.8
3	85	25.6	162	48.8	81	24.4	4	1.2
4	159	47.9	77	23.2	85	26.8	7	2.1
5	65	19.0	19	5.7	243	73.2	7	2.1
6	75	22.0	59	17.8	198	59.6	---	---
7	61	18.4	149	44.9	112	33.7	10	3.0
8	33	9.9	210	63.3	86	25.9	3	0.9
9	184	55.4	60	18.1	80	24.1	8	2.4
10	20	6.0	83	25.0	217	65.4	12	3.6
11	100	30.1	95	28.6	125	37.7	12	3.6
12	128	38.6	109	32.8	93	28.0	2	0.6
13	94	28.3	145	43.7	81	24.4	12	3.6
14	16	4.8	107	32.2	206	62.0	3	0.9
15	173	52.1	128	38.6	28	8.4	3	0.9
16	19	5.7	122	36.7	186	56.0	5	1.5
17	128	38.6	74	22.3	119	35.8	11	3.3
18	104	31.3	83	25.0	137	41.3	8	2.4
19	71	21.4	177	53.3	81	24.4	3	0.9

20	107	32.2	98	29.5	124	37.3	3	0.9
21	211	63.6	68	20.5	51	11.4	2	0.6
22	190	57.2	77	23.2	55	16.6	10	3.0
23	165	49.7	93	28.0	65	19.6	9	2.7
24	106	31.9	141	42.5	77	23.2	8	2.4
25	33	9.9	148	44.6	141	42.5	10	3.0
26	23	6.9	233	70.2	72	21.7	4	1.2
27	126	38.0	33	9.9	167	50.3	6	1.8
28	167	50.3	17	5.1	144	43.4	4	11.2
29	226	68.1	70	21.1	24	7.2	12	3.6
30	112	33.7	59	17.8	144	43.4	17	5.1

Nota. El resaltado indica la respuesta correcta.

Tabla 5. Frecuencia de distribución en los distractores por reactivo

Homogénea	Heterogénea	Frecuencia menor de 10%
11	18	12

Para un análisis completo que posibilite calibrar la prueba y hacerla válida es necesario revisar los valores de estos indicadores y en cada uno de los reactivos. Así como la discusión sobre la tabla de especificaciones en la que se establecen los temas, contenidos y resultados de aprendizaje que se pretende evaluar.

CONCLUSIONES

Por sus características, el presente estudio no pretende ser conclusivo, sin embargo, con base en los hallazgos es factible elaborar algunas consideraciones para continuar con esta tarea y estar en condiciones de hacer un estudio de seguimiento de los aprendizajes adquiridos en los distintos cursos de la LC.

Un primer aspecto a mencionar es que el colegio de profesores participante en este estudio, considera necesaria la evaluación continua de los aprendizajes obtenidos por los estudiantes para tomar las decisiones necesarias para la mejora de los programas y contenidos propuestos en las asignaturas. Asimismo, se pueden derivar nuevas propuestas didácticas que promuevan el aprendizaje significativo entre los estudiantes.

Contar con un instrumento con validez y bien calibrado será el inicio de lo que se convertirá en la posibilidad de conocer cuáles son los contenidos que mejor retienen los estudiantes y que



ISSN: 2448-6574

serán de relevancia para ellos en los cursos subsecuentes. Del mismo modo, se tendrá información base para instrumentar procedimientos de evaluación que se oriente al estudio de los procesos y con una aproximación global como propone Mateo (2005).

Como se mencionó antes el presente es el informe del estudio piloto asociado con el diseño de una prueba objetiva. En ese sentido, no es posible dar cuenta de los aprendizajes de los contenidos en los estudiantes. Hay diversas razones para evitar la tendencia de calificar el desempeño de los alumnos en la prueba. La primera es que los resultados de un piloto no deben utilizarse para evaluar el aprendizaje (CENEVAL, 2011). Además, la aplicación se realizó sin previa notificación a los participantes; esto se liga con que a los estudiantes se les dijo que su participación no tendría efecto alguno en su calificación.

Se considera esencial la presentación de los resultados al colegio de la Línea curricular. Con ello se pretende que una mayor cantidad de profesores se involucren en la tarea de discusión que enriquezca de manera significativa el análisis de los hallazgos, el análisis de los reactivos y su adecuación para alcanzar un instrumento válido y bien calibrado.

Una actividad más será la revisión y actualización de la tabla de especificaciones. Es importante analizar a la luz del propósito de la línea curricular y su aportación al perfil de egreso del psicólogo educativo, el programa del curso y los hallazgos aquí mostrados, a fin de identificar los aprendizajes significativos que los estudiantes deben poseer una vez aprobada la asignatura. De este modo, se abre la posibilidad de incrementar los contenidos evaluados y, por ende, la cantidad de reactivos que constituirán el instrumento en versiones posteriores.

REFERENCIAS

Backhoff, E., Larrazolo, N. y Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2(1). Consultado el 20 de junio de 2014 en: <http://redie.uabc.mx/index.php/redie/article/viewFile/15/26>



ISSN: 2448-6574

- Baghaei, P. y Amrahi, N. (2011). The effects of the number of options on the psychometric characteristics of multiple choice items. *Psychological Test and Measurement Modeling*, 53(2), 192-211.
- Bruno, J. y Dirkwager A. (1995). Determining the optimal number of alternatives to a multiple-choice test item: an information theoretic perspective. *Educational and Psychological Measurement*, 55(6), 959-966.
- CENEVAL (2000). Instructivo para la elaboración de reactivos. EGEL Pedagogía-Ciencias de la educación. México: CENEVAL.
- CENEVAL. (2011). Manual de elaboración de reactivos. México: CENEVAL.
- Haladyna, T. y Downing, S. (1993). How many options is enough for a multiple-choice test item? *Educational and Psychological Measurement*, 53, 999-1010.
- Haladyna, T.; Downing, S. y Rodriguez, M. (2002). A review of multiple-choice item-writing for classroom assessment. *Applied Measurement in Education*, 15(3), 309-334.
- Hernández, J., Pérez, C. G., Martínez, G., Bollás, P. & Dzib, A. (Julio de 2009). Plan de Estudios de la Licenciatura en Psicología Educativa, Plan 2009, Documento interno del programa de Psicología Educativa de la Universidad Pedagógica Nacional.
- Licenciatura en psicología educativa. (2011). Propuesta de evaluación de los cursos del plan 2009. Documento interno de trabajo del programa educativo en psicología educativa. Universidad Pedagógica Nacional.
- Lukas y Santiago (2009). *Evaluación educativa*. Madrid: Alianza.
- Mateo, J. (2005). *La evaluación educativa, su práctica y otras metáforas*. México: Alfaomega.
- Moreno, R.; Martínez, R. y Muñiz (2004). Directrices para la construcción de ítems de elección múltiple. *Psicothema*, 16(3), 490-407.
- Pérez L.C.G., Bollás, G.P., González, M.M.I, Álvarez, M.A.Y. y Castro, G.R. (2014). Examen general de conocimientos para la licenciatura en psicología educativa, plan 09,



ISSN: 2448-6574

Universidad Pedagógica Nacional. Indicadores psicométricos. Memorias Congreso Internacional de Educación, Evaluación.

Rodriguez, M. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, Summer, 3-13.

Scriven, M. (1967). The methodology of evaluation. En R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation*, 39-83. Chicago, IL: Rand McNally.