



EVALUACIÓN DE LOS APRENDIZAJES EN LA LICENCIATURA EN PSICOLOGÍA EDUCATIVA. DISEÑO DE UNA PRUEBA OBJETIVA PARA EL CURSO DE PSICOLOGÍA EVOLUTIVA DE LA ADULTEZ Y VEJEZ

MARÍA IMELDA GONZÁLEZ MECALCO

UNIVERSIDAD PEDAGÓGICA NACIONAL
mgonzalez@upn.mx

ALBA YANALTE ÁLVAREZ MEJÍA

UNIVERSIDAD PEDAGÓGICA NACIONAL
yalvarez@upn.mx

ROSA MARÍA NASHIKI ANGULO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
nashiki2001@yahoo.com.mx

RESUMEN

El trabajo tiene el propósito de diseñar una prueba objetiva para la evaluación de aprendizajes del curso Psicología evolutiva de la adultez y vejez de tercer semestre de la Licenciatura en Psicología Educativa. Se construyó una prueba de 41 reactivos elaborados a partir de los contenidos del curso. Esta fue aplicada a una muestra de estudiantes, cuyos resultados permitieron obtener el índice de facilidad y discriminación de los reactivos a fin de determinar la funcionalidad de los mismos.

Se muestra un reporte parcial que pretende evaluar los aprendizajes de los estudiantes desde una perspectiva formativa, por lo que se presentan resultados solo del diseño de la prueba objetiva, que permita el avance, discusión y formulación de una evaluación que considere el diseño de los programas, la experiencia de los estudiantes y la didáctica de los docentes.

Palabras clave: Evaluación del aprendizaje, Pruebas de aprovechamiento, Validez.





INTRODUCCIÓN

La Licenciatura en Psicología Educativa (LPE) en la Universidad Pedagógica Nacional se cursa a lo largo de ocho semestres, organizados en tres Fases y siete Líneas Curriculares. Entre estas se encuentra *Desarrollo y Aprendizaje*, cuyo 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 Línea se compone por siete asignaturas, de las cuales tres abordan el estudio del desarrollo y cuatro el área del aprendizaje. Las primeras se enfocan al estudio del desarrollo evolutivo de las personas: infancia, adolescencia, adultez y vejez. En este último, se pretende generar un instrumento para la evaluación de los aprendizajes, por lo que el presente documento es el resultado del estudio piloto del instrumento elaborado.

El estudio es parte de un trabajo de investigación más amplio (que incluye los cursos de *Psicología evolutiva de la infancia* y *Psicología evolutiva de la adolescencia*), en esta primera etapa del trabajo, se tomó una muestra de estudiantes de cuarto semestre de la LPE, quienes cursaron la materia evaluada en tercer 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 adultez y vejez, incluidos en el programa de estudio.

REFERENTES TEÓRICOS

La evaluación educativa puede ser vista desde distintas perspectivas teóricas, mismas que le asignan funciones específicas. Ketele y Roegiers (citado en Lukas y Santiago, 2009) hacen referencia a funciones tales como regulación, predicción, descripción, entre otras.

Desde hace varios años, 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.

Este trabajo pretende acercarse a una evaluación de carácter formativo, siendo coincidente con el planteamiento de Mateo (2005) cuando afirma que dependiendo del momento o de la necesidad del proceso de evaluación, se opta por una u otra función de la evaluación. De tal manera que la investigación de este reporte parcial informa de un acercamiento a una evaluación formativa, sin embargo, en esta primera etapa se da cuenta solamente del comportamiento de los reactivos de la prueba objetiva.

Para llevar a cabo la evaluación de los aprendizajes del curso mencionado, se revisaron diversos procedimientos para la recogida de información. Lukas y Santiago (2009) presentan una clasificación que da cuenta de aquellas técnicas de corte cuantitativo y aquellas que se dirigen a una obtención de información más bien cualitativa. Este trabajo pretende acercarse a una evaluación de corte cualitativo, parte de la aplicación de una prueba objetiva para dar un panorama del aprendizaje sobre los contenidos del curso psicología evolutiva de la adultez y vejez por los estudiantes de la LPE.

Con base en que las pruebas diseñadas para evaluar los aprendizajes de los estudiantes deben ser estandarizadas y validadas, no solo por el impacto que algunas de estas tienen sobre los resultados de los estudiantes, sino porque la toma de decisiones debe estar fundamentada, se consideró la facilidad/dificultad del reactivo y la discriminación del mismo como centro de análisis. La dificultad entendida como la proporción de personas que responden correctamente a un reactivo, de tal suerte que entre mayor sea la proporción, menor es su dificultad. Con la discriminación se espera que aquellos con una puntuación alta en toda la prueba tengan más probabilidades de contestar correctamente al reactivo y que ocurra a la inversa con el estudiante que obtuvo puntuaciones bajas (Backhoff, Larrazolo y Rosas, 2000).

Se considera un puntaje entre 0.28 y 0.72 como un índice de facilidad apropiado, superior a esto se considera un reactivo que la mayoría responde como correcto, por lo que se sugiere revisar las opciones y la estructura del reactivo para decrementar su nivel de facilidad. Por el contrario un reactivo con un índice inferior a 0.28 tiende a provocar que pocos sustentantes lo respondan correctamente por lo que también será necesario revisar las opciones de respuesta,





ya que existe la posibilidad de que alguna de ellas sea parcialmente correcta o que su estructura sea confusa para el sustentante, con ello se pretende incrementar su nivel de facilidad.

A fin de mejorar el instrumento de evaluación, se retoman las recomendaciones de Ebel y Frisbie (citado en Backhoff, Larrazolo y Rosas, 2000) derivadas del puntaje de discriminación de los reactivos: un índice de discriminación superior a 0.39 se considera un reactivo de excelente calidad, de 0.30 a 0.39 buena, de 0.20 a 0.29 regular, de 0.00 a 0.19 pobre, menor a 0.00 pésima. La recomendación indica que los reactivos de excelente calidad se conservan, los de buena se pueden mejorar, los regulares se revisan, los pobres se descartan o revisan a profundidad y los pésimos se descartan.

MÉTODO

Participantes

Se trabajó con una muestra de 258 estudiantes de cuarto semestre de la carrera en Psicología Educativa, de los cuales 131 pertenecen al turno matutino y 127 al vespertino. Todos eran estudiantes inscritos en la carrera al momento de responder la prueba y habían cursado el semestre previo la materia Psicología Evolutiva de la Adultez y Vejez.

Escenario

El estudio se llevó a cabo en la Universidad Pedagógica Nacional, Unidad Ajusco. El instrumento se aplicó en las aulas de clases, en horario escolar y de manera simultánea.

Instrumento

Se diseñó una prueba objetiva a partir de la revisión del *Programa del curso* y del *Documento técnico para la elaboración de los reactivos del Examen General de Conocimientos*. Del primero se retoma la organización de los contenidos (propósito, unidades, temas) y del segundo la estructura jerárquica de los temas y conceptos fundamentales, mismos que sustentan el diseño del Examen de conocimientos para la titulación (Pérez, Bollás, González, Álvarez y Castro, 2014).

Se tomaron en cuenta los criterios que proponen Haldayna, Downing y Rodríguez (2002) y Moreno, Martínez y Muñiz (2004), en relación a que los reactivos serían contruidos con tres opciones de respuesta. Lo anterior fue decidido a partir de los informes de Baghaei y Amrahi (2011), Bruno y Dirkzwager (1995), Haldayna y Downing (1993) y Rodríguez (2005), en los que





se menciona que en los exámenes de elección múltiple, tres opciones distractoras son suficientes. La parte medular en el diseño de los reactivos no radica en el número de distractores, sino en la calidad de los mismos.

En los trabajos referidos se menciona que en este tipo de exámenes, pensar y decidir cuál será la cuarta opción es una tarea compleja, porque es necesario seleccionar un concepto o construir una frase perteneciente al mismo campo semántico y, al mismo tiempo, que sea para el sustentante una opción suficientemente factible y posible. En consecuencia, advierten los autores citados, la última opción resulta para el participante una situación obviamente falsa, lo cual se demuestra con la reducida frecuencia de elección.

Con base en el programa del curso se elaboró una primera lista de temas. Posteriormente, se procedió a realizar los reactivos de acuerdo con la importancia de los temas para la formación del psicólogo educativo, tanto a nivel del curso en particular, como a nivel transversal, ya que son muy importantes para otros cursos en esta carrera.

Una vez que el grupo de docentes participantes en este trabajo revisó y estuvo de acuerdo en los aprendizajes a evaluar a través de los reactivos diseñados, la prueba quedó con 41 reactivos, organizados por unidad como se muestra en el tabla 1.

PROCEDIMIENTO

El presente trabajo está organizado en tres fases. La primera está centrada en analizar el contenido del curso y elaborar la prueba objetiva. Para ello, se invitó a docentes que habían impartido el curso al menos una vez a participar en el diseño del examen. De manera colegiada, cada profesor compartió los reactivos diseñados para las diversas unidades del curso, se discutió su relevancia, relación con los contenidos y propósitos establecidos en el programa de la materia. Una vez revisados los reactivos, se organizaron en la prueba objetiva y se realizó la edición final de la misma.

La fase dos consistió en la aplicación de la prueba, lo cual implicó solicitar el apoyo a los docentes de 4° semestre para la aplicación a sus estudiantes de manera simultánea en los grupos. La fase tres consistió en el análisis de la información mediante una base de datos y, un análisis de discriminación y dificultad de cada reactivo, dicha información se entregará a los





participantes para su discusión en colegiado y toma de acuerdos sobre acciones a realizar a partir de los datos obtenidos.

RESULTADOS

De los resultados de la prueba se obtuvo el índice de facilidad/dificultad y el de discriminación por cada uno de los reactivos. De acuerdo con estos datos se encuentra que ocho reactivos tienen un índice de discriminación de excelente y buena calidad, es decir, pueden ser considerados reactivos que diferencian a quienes saben la respuesta de quienes no la saben y eligen al azar. De nivel de discriminación regular hay 18 reactivos y de calidad pobre hay 15 reactivos (uno obtiene una discriminación negativa), para estos es recomendable hacer una revisión de su estructura así como de sus distractores (véase tabla 2).

De acuerdo con los puntajes de facilidad se puede observar en la tabla 2 que se cuenta con cinco reactivos (subrayados con línea punteada) en los que su nivel de facilidad es bajo y seis reactivos (subrayados con doble línea) en los que su nivel de facilidad es alto; para ambos casos se recomienda hacer una revisión de la estructura del reactivo, así como del comportamiento de sus opciones de respuesta para identificar qué aspecto es oportuno modificar e identificar mejor el aprendizaje de los estudiantes.

Se realizó un análisis del comportamiento de las opciones de respuesta de todos los reactivos, a continuación únicamente se muestran los resultados de los reactivos cuyo índice de dificultad o de discriminación se encuentra fuera de lo óptimo.

La tabla 3 muestra que en los cinco reactivos en línea punteada la opción correcta es demasiado evidente en comparación con el resto de las opciones, por lo que la mayoría de los sustentantes la eligen como la correcta (las negrillas en la tabla indican la respuesta correcta).

Para los nueve reactivos subrayados es necesario revisar la estructura de la base y los distractores para identificar qué es lo que hay que modificar, ya que una de las opciones de respuesta pareciera ser la correcta, cuando en realidad es errónea.

En los reactivos r5, r6, r19 y r36 se comportan apropiadamente las opciones de respuesta, por lo que la sugerencia es revisar el contenido y estructura general.





En términos de resultados del examen, puede comentarse que la calificación más alta que se obtuvo fue de 78.05 y la más baja 29.27. Se considera que los puntajes son bajos para ser un curso aprobado el semestre inmediato anterior. Estos resultados no muestran una distribución normal, debido en parte a que es la primera vez que se aplica el instrumento y se tiene el mínimo del muestreo requerido, por lo tanto, no es posible tener contrastes entre generaciones ni relación con las calificaciones de los estudiantes en este curso.

A la par de la evaluación de los aprendizajes, se realizó un grupo de discusión con los docentes que han impartido el curso para evaluar el programa. Producto de esta discusión, se reconoció la necesidad de incorporar materiales y temas centrados en la etapa de la adultez, ya que sobre esta hay menos información en el programa; de igual manera se hace necesario revisar los contenidos en un continuo del desarrollo que abarque ambas etapas evolutivas.

Se identificaron temas importantes que deben explicitarse en el programa del curso, ya que se revisan y evalúan con resultados favorables. Sin embargo, es necesario tomar en cuenta los resultados obtenidos para actualizar el programa del curso.

CONCLUSIONES

El trabajo permite concluir lo siguiente:

- Los criterios de discriminación y facilidad permitieron identificar de manera efectiva el comportamiento de los reactivos en particular y del examen en general. Con ello, se da pauta para la creación de instrumentos con un grado de validez aceptable, al reconocer los aspectos de mejora.
- Los resultados a partir de los índices de discriminación, de facilidad y del comportamiento de las opciones de respuesta, permiten determinar el tipo de decisión que se tomará con cada uno de los reactivos, de tal manera que estos puedan seguir probándose en futuras aplicaciones y con ello, mejorar el instrumento.
- La aplicación de este instrumento, a una muestra y el análisis de su comportamiento dan pie al desarrollo de instrumentos similares con otros cursos de la Línea Curricular Desarrollo y Aprendizaje, de tal manera que a largo plazo pueda contarse con instrumentos probados para evaluar los aprendizajes que se construyen en los cursos de esta línea.





- La construcción de este tipo de instrumentos permite avanzar hacia la evaluación formativa y objetiva que dé cuenta de los aprendizajes de los estudiantes de psicología educativa.
- Para continuar con este tipo de evaluaciones es necesario realizar un análisis detallado del contenido de los reactivos y de su funcionamiento, para garantizar que el examen incluya los contenidos esenciales y se muestre equilibrado en función de la currícula.
- Terminado el análisis de los datos, sigue la discusión con el equipo de docentes para tomar decisiones respecto a mejoras en el programa y en la prueba misma.

TABLAS

Tabla 1. Temas y conceptos de Psicología Evolutiva de la Adultez y Vejez

Unidad	Tema	Concepto
	Precurrentes	Desarrollo, cambio y aprendizaje
1	Edad y relación con adultez y vejez	Edad Yo intemporal
	Modelos explicativos	Mecanicista, contextual, Ciclo vital, Ecológico
	Influencia de factores socioculturales	Influencia no normativa
2	Cambio en los sistemas corporales	Capacidad de reserva
	Enfermedades, prevención y mantenimiento de la salud	Estilo de vida Enfermedad y malestar Salud
3	Contexto familiar (estructura, matrimonio, divorcio, viudedad, casarse de nuevo, rol del abuelo)	Crisis de la mitad de la vida Divorcio Estructura familiar Rol del abuelo Duelo, pena, luto
	Contexto laboral	Jubilación
4	Memoria y atención	Memoria
	Procesos superiores y aprendizaje (inteligencia, creatividad y sabiduría)	Inteligencia Sabiduría





Tabla 2. Índice de facilidad y discriminación

Reactivo	Facilidad	Discriminación
<u>1</u>	<u>0.95</u>	0.06
<u>2</u>	<u>0.27</u>	0.21
3	0.59	0.37
4	0.50	0.24
5	0.55	0.19
6	0.62	0.17
7	0.28	0.30
8	0.44	0.20
9	0.45	0.27
10	0.52	0.34
11	0.61	0.20
<u>12</u>	<u>0.91</u>	0.11
<u>13</u>	<u>0.84</u>	0.20
14	0.71	0.21
15	0.52	0.39
<u>16</u>	<u>0.84</u>	0.13
17	0.34	0.23
<u>18</u>	<u>0.16</u>	0.16
19	0.49	-0.07
20	0.70	0.29
21	0.36	0.33

Reactivo	Facilidad	Discriminación
22	0.52	0.29
23	0.47	0.31
24	0.41	0.23
<u>25</u>	<u>0.17</u>	0.00
26	0.35	0.21
<u>27</u>	<u>0.22</u>	0.13
28	0.33	0.14
29	0.28	0.11
30	0.72	0.29
31	0.56	0.27
32	0.34	0.06
<u>33</u>	<u>0.21</u>	0.26
34	0.48	0.17
35	0.45	0.24
36	0.53	0.14
37	0.34	0.30
<u>38</u>	<u>0.79</u>	0.21
39	0.27	0.27
<u>40</u>	<u>0.85</u>	0.19
41	0.55	0.34





Tabla 3. Selección de opciones de respuesta

Reactivo	Porcentaje de respuesta elegida			
	No contestó	A	B	C
r1	0	0	4.7	<u>95.3</u>
r2	0.4	2.3	<u>27.1</u>	70.2
r5	0	29.5	<u>54.7</u>	15.9
r6	0	62	11.6	26.4
r12	0	2.7	<u>90.7</u>	6.6
r16	0.8	2.3	12.8	<u>84.4</u>
r18	0.4	67.4	16.7	<u>15.5</u>
r19	0	30.6	<u>49.2</u>	20.2
r25	0	51.2	<u>17.4</u>	31.4
r27	0.4	<u>22.5</u>	8.9	68.2
r28	0	53.1	14.3	<u>32.6</u>
r29	0.8	63.6	<u>27.9</u>	7.8
r32	0.8	34.1	<u>34.1</u>	31.0
r33	0.4	57.8	<u>20.5</u>	21.3
r34	0.4	<u>47.7</u>	7.8	44.2
r36	0	11.6	35.7	<u>52.7</u>
r39	0	6.6	<u>78.7</u>	14.7
r40	0.8	7.8	<u>85.3</u>	6.2

BIBLIOGRAFÍA Y 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>
- 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 Dirkzwager 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.
- 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 Rodríguez, M. (2002). A review of multiple-choice item-writing for classroom assessment. *Applied Measurement in Education*, 15(3), 309-334.
- 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, J. (2004). Directrices para la construcción de ítems de elección múltiple. *Psicothema* 16(3), 490-497. Consultado el 11 de marzo de 2014 en: <http://www.psicothema.com/pdf/3023.pdf>
- Pérez, C., Bollás, P., González, M., Álvarez, A. y Castro, R. (2014). Examen general de conocimientos para la licenciatura en psicología educativa, plan 09, Universidad Pedagógica Nacional. Indicadores psicométricos. *Memorias Congreso Internacional de Educación: Evaluación*.
- Rodríguez, 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.

