

Número Publicado el 26 de Agosto de 2017

<http://dx.doi.org/10.23857/dom.cien.pocaip.2017.3.5.agos.374-389>

URL:<http://dominiodelasciencias.com/ojs/index.php/es/index>

Ciencias Pedagógicas

Artículo Científico

Conocimiento pedagógico en evaluación educativa de los docentes y la construcción de exámenes del idioma inglés

Pedagogical knowledge in educational evaluation of teachers and construction of English language exams

Amilcar O. Gonzalez-Diaz ^I
Universidad Central del Ecuador
aogonzalez@uce.edu.ec

Edison S. Sanguña-Loachamin ^{II}
Universidad Central del Ecuador
essanguna@uce.edu.ec

Gabriela M. Moso-Mena ^{III}
Universidad Central del Ecuador
gmmoso@uce.edu.ec

Boris E. Vasconez-Torres ^{IV}
Universidad Central del Ecuador
bevasconez@uce.edu.ec

Recibido: 30 de enero de 2017 * **Corregido:** 20 de febrero de 2017 * **Aceptado:** 20 junio de 2017

- ^{I.} Magister en Lingüística y Didáctica de la Enseñanza de Idiomas Extranjeros; Diploma Superior de la Enseñanza de Inglés como Segunda Lengua; Licenciado en Ciencias de la Educación mención Inglés; Técnico Superior en Turismo; Universidad Central del Ecuador.
- ^{II.} Magister en Docencia Universitaria y Administración Educativa; Especialista en Diseño Curricular por Competencias; Licenciado en Ciencias de la Educación mención: Plurilingüe; Universidad Central del Ecuador.
- ^{III.} Magister en Educación Superior y Equidad de Género; Licenciada en Ciencias de la Educación mención Plurilingüe; Universidad Central del Ecuador.
- ^{IV.} Magister en Educación; Especialista en Currículo y Didáctica; Diploma Superior en Investigación Socioeducativa; Ingeniero en Diseño Gráfico Empresarial; Licenciado en Lingüística Aplicada al Idioma Inglés; Licenciado en Ciencias de la Educación mención Diseño y Dibujo; Universidad Central del Ecuador.

Resumen

El propósito de este estudio es verificar la relación existente entre el conocimiento en evaluación educativa y la producción de exámenes de logro, llamados en inglés achievement tests. La primera etapa del trabajo constituyó en la revisión literaria del tema y luego se procedió a diseñar dos instrumentos de recolección de datos, una encuesta encaminada a medir el conocimiento en evaluación aplicada a docentes, y una rúbrica para verificar la calidad de las pruebas construidas en ese entonces. Después, los datos obtenidos en ambos instrumentos fueron analizados y contrastados encontrando que, mientras la encuesta aplicada muestra un alto conocimiento en evaluación, la calidad de exámenes difiere de aquel conocimiento. Se concluyó en la inexistencia de la relación entre conocimiento y calidad de pruebas en este grupo de docentes. Finalmente, se propone una Guía como una herramienta que ayude a mejorar la calidad de pruebas.

Palabras clave: Evaluación educativa; exámenes de logro; pruebas estandarizadas; diseño de pruebas; tipos de reactivos.

Abstract

The purpose of this study is to find the existing relationship between knowledge in educational assessment and achievement test production. The first phase went through literature review of both variables, and then two instruments to collect data were designed. A survey aiming at measuring teachers' knowledge of evaluation and assessment, and a rubric which aimed at verifying the quality of tests already collected. After that, the data obtained in both instruments were analyzed and contrasted finding that teachers have a high knowledge of assessment and evaluation, but that knowledge differs from test quality. Finally, a Guide is proposed as a tool that helps future test quality.

Key words: Educational evaluation; achievement tests; standardized tests; test design; item types.

Introducción.

La evaluación educativa ha sido un tema de discusión por muchos estudiosos desde tiempos antiguos. En Ecuador los procesos evaluativos han sufrido varios cambios en los últimos tiempos. Es importante recordar que actualmente en el Ecuador existen varios organismos encargados de la evaluación educativa, entre ellos, el Sistema Nacional de Evaluación, el Instituto Nacional de Evaluación Educativa, el Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior, etc. Estas instituciones y otras son las encargadas de regir los destinos de la evaluación educativa en nuestro país. En este contexto, la evaluación en el Ecuador está ganando mayor importancia día tras día. (González Melo & Ospina Serna, 2013) (Bermúdez-Jaimes & Mendoza-Páez, 2008)

La presente investigación tiene por objetivo principal denotar la relación existente entre lo que conocen los docentes en evaluación con la producción de exámenes de logro que estos producen. La relevancia de este trabajo recae en el tipo de pruebas que se diseñan ya que de ello depende la promoción o no de un estudiante. Es necesario mencionar que la evaluación se divide principalmente en dos, una formativa y otra sumativa. La primera encargada de los procesos y la otra en el resultado. Debe existir una estrecha relación entre estos dos tipos de evaluación. Los exámenes de logro o achievement tests en inglés, están presentes durante todo el proceso de la evaluación. Las pruebas de lección o de unidad son consideradas también pruebas de logro. (Vasco, 1990) (Careaga, 2001)

Los resultados de este trabajo investigativo no serán únicamente relevantes para la Universidad Central del Ecuador sino también para cualquier institución que viva contextos similares. Alrededor de todo el país, las instituciones educativas están sufriendo una fase de

transformación donde la eficiencia es su común denominador. Así, la construcción de pruebas, estrechamente relacionada con la evaluación se constituye en un pilar que asegure la calidad educativa. (Montenegro, 2003)

La presente investigación tiene como objetivo analizar la relación existente entre el conocimiento pedagógico en evaluación de los profesores y la construcción de exámenes en el Centro Universitario de Idiomas de la Universidad Central del Ecuador.

Materiales y métodos.

Al diseñar la investigación, se debe tomar en cuenta la importancia de la misma ya que constituye la columna vertebral del estudio. De igual forma, al seguir los protocolos científicos de investigación, brindamos mayor validez al trabajo propuesto. Con éste preámbulo, a continuación se hace mención de la estructura holística del presente trabajo de investigación, tomando en cuenta el tipo de investigación, diseño de la investigación, y modalidad de la investigación.

La confiabilidad y validez de los resultados de este trabajo investigativo depende en gran medida de la obtención de los resultados, los cuales a su vez deben ser igualmente confiables y válidos.

Para Yuni y Urbano, (2006) la función primaria de las técnicas de recolección de datos es la observación y el registro de los fenómenos. Para esta investigación, se ha seleccionado la encuesta y observación directa como técnicas de recolección de información primaria. Recordemos que el tipo de técnica que se seleccione y los instrumentos de recolección, responden a los objetivos de la investigación y a la forma de plantear el problema (Rojas, 2001). El objetivo principal de este trabajo es encontrar la relación existente entre el conocimiento pedagógico en evaluación y la

Conocimiento pedagógico en evaluación educativa de los docentes y la construcción de exámenes del idioma inglés

construcción de exámenes y se planteó como problema de investigación descubrir si este conocimiento afecta o no la construcción de pruebas. Por ende se ha seleccionado a la encuesta como la técnica idónea para recolectar información de los docentes y la observación directa como la otra técnica de recolección para analizar los exámenes de logro producidos por los maestros. (Vidal, 2001) menciona que la encuesta es una técnica ventajosa debido a que cubre un área extensa logrando así obtener información de muchos individuos. Además, los encuestados tienen una sensación de anonimato, lo que permite que los datos sean fiables.

Resultados.

Diseño de pruebas

Cuadro N° 1.- Diseño de pruebas - Frecuencias

INDICADOR	FRECUENCIA	%
EXCELENTE	2	2,22
SATISFACTORIO	10	11,11
PUEDA MEJORAR	50	55,56
INADECUADO	28	31,11
TOTAL	90	100,00

Cuadro N° 2.- Diseño de pruebas - Rúbrica

Criterios	Nivel			
	4. Excelente	3. Satisfactorio	2. Puede mejorar	1. Inadecuado
Diseño de pruebas	La prueba contiene objetivos precisos, y instrucciones claras que no confunden al estudiante.	La prueba contiene objetivos y especificaciones que no confunden al estudiante.	La prueba contiene objetivos y instrucciones que confunden al estudiante.	La prueba no contiene objetivos pero si contiene instrucciones que no son claras.

Análisis en interpretación

Se observa que la mayoría de las pruebas contienen instrucciones que no son claras y se prestan a doble interpretación. Recordemos que las instrucciones deben ser precisas, e indiquen lo que debe realizar el estudiante. Se observa que no existen objetivos en una gran cantidad de pruebas. Los objetivos de la prueba indican al estudiante el propósito que tiene el docente para utilizar dicho instrumentos. Es posible que el docente únicamente necesite diagnosticar o evaluar parte del contenido. Estos dos aspectos de: objetivos e instrucciones son sumamente importantes ya que los educandos sabrán qué y cómo responder. Finalmente se observa un bajo número de exámenes que no contienen instrucciones claras, por ello es indispensable que los docentes cumplan con las necesidades académicas requeridas.

Pruebas de las habilidades comunicativas

Cuadro N° 3.- Pruebas de las habilidades - Frecuencias

INDICADOR	FRECUENCIA	%
EXCELENTE	0	0,00
SATISFACTORIO	18	20,00
PUEDA MEJORAR	45	50,00
INADECUADO	27	30,00
TOTAL	90	100,00

Cuadro N° 4.- Pruebas de las habilidades comunicativas - Rúbrica

Criterios	Nivel			
	4. Excelente	3. Satisfactorio	2. Puede mejorar	1. Inadecuado
Pruebas de habilidades comunicativas	La prueba contiene las cuatro habilidades: lectura, escritura, comprensión auditiva, producción oral.	La prueba contiene tres habilidades del idioma.	La prueba contiene dos habilidades del idioma.	La prueba contiene solo una habilidad del idioma.

Análisis en interpretación

Una prueba debe estar estructurada con las habilidades y micro habilidades de un idioma, sin embargo, se observa que la mayoría únicamente cuenta con una o dos habilidades que se repiten entre escritura gramática/ vocabulario y lectura. Estos resultados contradicen todos los estudios realizados por Brown (2010) quien manifiesta que una prueba debería estar balanceada de forma que se mida la mayoría de las competencias de los estudiantes.

Pruebas de las micro-habilidades

Cuadro N° 5.- Pruebas de las micro-habilidades - Frecuencias

INDICADOR	FRECUENCIA	%
EXCELENTE	0	0,00
SATISFACTORIO	54	60,00
PUEDE MEJORAR	31	34,44
INADECUADO	5	5,56
TOTAL	90	100,00

Cuadro N°6.- Pruebas de las micro-habilidades - Rúbrica

Criterios	Nivel			
	4. Excelente	3. Satisfactorio	2. Puede mejorar	1. Inadecuado
Pruebas de las micro-habilidades	Aparte de las habilidades del idioma, la prueba contiene vocabulario y gramática claramente balanceado su porcentaje.	Aparte de las habilidades del idioma, la prueba contiene vocabulario y gramática pero no está balanceado su porcentaje.	Aparte de las habilidades del idioma, la prueba contiene únicamente gramática.	Aparte de las habilidades del idioma, la prueba contiene únicamente vocabulario

Análisis en interpretación

Como se manifestó en el ítem anterior las pruebas en su mayoría contienen entre una o dos habilidades del idioma, pero en su mayoría contiene gramática y vocabulario. Desafortunadamente, muchas pruebas contienen solo gramática o solo vocabulario, contradiciendo el hecho de que una prueba debe estar distribuida de tal forma los estudiantes sean evaluados en sus múltiples competencias. Idealmente una prueba bien elaborada debe evidenciar cuanto realmente conoce un estudiante del idioma, para ello se debe elaborar instrumentos que recojan mayoritariamente las habilidades. Se requiere que los docentes conozcan la estructura de una prueba de tal manera que su construcción se facilite y se alcance verdaderamente a evaluar los aprendizajes.

Agrupación de preguntas y formatos.

Cuadro N° 7.- Pruebas elaboradas por los docentes - Frecuencias

INDICADOR	FRECUENCIA	%
EXCELENTE	12	13,33
SATISFACTORIO	9	10,00
PUEDE MEJORAR	63	70,00
INADECUADO	6	6,67
TOTAL	90	100,00

Cuadro N° 8.- Agrupación de preguntas y formatos - Rúbrica

Criterios	Nivel			
	4. Excelente	3. Satisfactorio	2. Puede mejorar	1. Inadecuado
Pruebas elaboradas por el profesor.	La agrupación y organización de las preguntas y el formato de la prueba son bien claras y precisas.	La agrupación y organización de las preguntas y el formato de la prueba son medianamente claras.	La agrupación y organización de las preguntas y el formato de la prueba son poco claras.	La agrupación y organización de las preguntas y el formato de la prueba no son claras.

Análisis en interpretación

Tomando como referencias a la literatura encontrada y expuesta en el marco teórico, las preguntas deben estar agrupadas por bloques, es decir por habilidades que se desea medir. Las preguntas deben estar numeradas desde el inicio hasta el final consecutivamente y además deben tener su valor cuantitativo. El tamaño de la letra y el espacio entre líneas deben ser amigables ocupando el ancho de la hoja. Al decir amigables indica legible que no implique esfuerzo ni para el estudiante, ni para el profesor. Comúnmente se empieza una prueba con la habilidad de comprensión

auditiva pero no es un estándar o regla. La prueba debe tener identificación, es decir la institución, el tipo de prueba, y objetivos, además de los datos del estudiante.

Se observa que las pruebas en su mayoría no cumplen con las especificaciones anteriormente mencionadas. Lamentablemente algunas pruebas únicamente contienen los datos del educando y las preguntas. Se observa además diferente tipo de letra, espacios, márgenes, y equivalencias de las preguntas. Todo esto contradice los elementos básicos de la evaluación, valoración y construcción de exámenes.

Tipos de reactivos

Cuadro N° 9.- Tipos de reactivos

INDICADOR	FRECUENCIA	%
EXCELENTE	12	13,33
SATISFACTORIO	33	36,67
PUEDE MEJORAR	4	4,44
INADECUADO	41	45,56
TOTAL	90	100,00

Cuadro N° 10.- Tipos de reactivos

Criterios	Nivel			
	4. Excelente	3. Satisfactorio	2. Puede mejorar	1. Inadecuado
Tipos de reactivos	La prueba cuenta con la combinación de reactivos de respuesta estructurada y reactivos de respuesta abierta.	La prueba cuenta con reactivos mayoritariamente de respuesta estructurada y pocos de respuesta abierta	La prueba cuenta con reactivos mayoritariamente de respuesta abierta y pocos de respuesta estructurada.	La prueba cuenta solo con uno de los tipos de reactivos.

Análisis en interpretación

Una prueba debe contener diferentes tipos de preguntas de forma que los estudiantes muestren sus conocimientos de forma holística. Se puede apreciar que mayoritariamente las pruebas contienen un solo tipo de reactivo. Este hecho lleva a los estudiantes a realizar una prueba mecánicamente o a desarrollar preguntas de respuesta abierta. Al hablar de forma mecánica, se hace referencia a las preguntas de opción múltiple. Idealmente, una prueba debe contener los dos tipos de reactivos: Opción múltiple y preguntas de respuestas abierta para el caso de la escritura, por ejemplo.

Discusión.

Es importante recordar que se aplicaron dos instrumentos de recolección de información, una encuesta, encaminada a determinar el conocimiento en evaluación y construcción de exámenes que poseen las/los docentes; y el otro instrumento una rúbrica, que tuvo por objetivo establecer la calidad de exámenes realizados por las/los docentes. Como se manifiesta anteriormente, se aplicó dos instrumentos, uno por cada variable de la investigación.

La encuesta reflejó resultados satisfactorios debido a los altos porcentajes obtenidos en cada ítem del cuestionario. Se encontró que los docentes conocen los principios de la evaluación como la practicidad, validez, confiabilidad, medición, diferencia entre evaluación y valoración. Así mismo, la encuesta refleja que los docentes conocen los procesos de construcción de pruebas y saben diferenciar entre una prueba de logro, de aptitud, estandarizada, de diagnóstico, y de proficiencia. De la misma manera, las/los docentes encuestados manifiestan estar al tanto en relación a la evaluación alterativa, los diarios, portafolios y rúbricas.

Como se declara en el párrafo anterior, los docentes expresan su alto conocimiento en construcción de pruebas, principios de la evaluación, tipos de pruebas y preguntas. Lamentablemente esta información contradice los resultados obtenidos en la Rúbrica aplicada a las pruebas realizadas por los mismos docentes.

En relación a la rúbrica, existen varios parámetros que fueron observados. En primer lugar se observó el diseño de las pruebas, su contenido, tipos de preguntas, y número de preguntas para cada habilidad, sin embargo, a pesar que los docentes manifiestan conocer el proceso de construcción de pruebas, esto no se observa en la Rúbrica. La rúbrica refleja bajos valores en el parámetro Diseño de pruebas.

Además, se observó si las pruebas miden las cuatro habilidades principales del idioma, pero solo un bajo porcentaje de la pruebas evalúa las cuatro habilidades. Mayoritariamente las pruebas miden gramática y vocabulario.

Finalmente, se puede manifestar que existe una contradicción entre los conocimientos en evaluación que posee las/los docentes y las pruebas. Se debería encontrar que mientras más conocimientos poseen las/los docentes, de mejor calidad serán las pruebas. Sin embargo, los dos instrumentos reflejan lo contrario. La encuesta evidencia un alto conocimiento en evaluación y sus procesos, pero la rúbrica arroja bajos porcentajes en la calidad de pruebas.

Conclusiones.

Después de haber realizado el análisis e interpretación de los datos obtenidos en los dos instrumentos de recolección de información: la encuesta y la Rúbrica, a continuación se enumeran las conclusiones del estudio:

En relación a la encuesta, (García, 2012) menciona que una de las desventajas es que “existe menor sinceridad por parte de los entrevistados” (p. 84). Este hecho es evidente en el presente trabajo investigativo ya que los resultados de la encuesta no coinciden con los resultados de la rúbrica. Mientras que la encuesta aplicada a los docentes indica un alto conocimiento de evaluación, la rúbrica, donde se observa los exámenes, revela todo lo contrario. Los exámenes muestran bajos niveles de preparación en evaluación educativa y principalmente de elaboración de pruebas.

Debido a que no existe una preparación previa en evaluación educativa y elaboración de pruebas específicamente, los docentes realizan exámenes sin una guía que les indique los parámetros que deben seguir para construir pruebas con calidad.

El alto porcentaje reflejado en la encuesta debería resultar en exámenes bien elaborados, sin embargo, es todo lo contrario. Se ha encontrado exámenes que no alcanzan un mínimo de estándares requeridos.

El texto utilizado para la enseñanza del idioma Inglés cuenta con las cuatro habilidades básicas, sin embargo la mayoría de los exámenes no miden estas cuatro destrezas, sino se limitan a evaluar mayoritariamente Gramática y vocabulario. Con esta realidad es fácil predecir que los exámenes no miden los conocimientos alcanzados en el aula de clase.

Los exámenes sometidos a la rúbrica, cuentan con distintos formatos entre ellos: de identificación de la prueba, número de preguntas, tipo de preguntas, orden que aparecen las habilidades medibles, tipo de letra, interlineado, valor a cada ítem, objetivos e instrucciones.

Recomendaciones.

En primer lugar, debería existir un departamento de evaluación con el fin de guiar el proceso evaluativo, entre ello: entrenamiento en evaluación educativa, elaboración de exámenes incluyendo pilotaje, elaboración de varias formas de exámenes, seguimiento de los resultados de las exámenes, administración y calificación de exámenes, calendarización de toma de exámenes, y manejo de exámenes para evitar que se filtren entre los estudiantes.

Así mismo, se debe elaborar objetivos claros de aprendizaje por cada nivel de forma que los exámenes respondan a dichos objetivos. Independientemente del texto utilizado, se debe establecer las políticas, misión, visión, objetivos, metas a alcanzar en un determinado periodo, así, cada instrumento de evaluación responderá a dicha planificación.

Realizar periódicamente cursos de actualización en evaluación educativa y construcción de exámenes. De no contar con capacitadores en el área, se debería elaborar una guía que indique los principios de la evaluación educativa y los pasos a seguir para construir exámenes con calidad.

Se debe evitar elaborar pruebas por equipos que a su vez distribuyen el trabajo a sus partes, de esta forma, los exámenes son el resultado de la unión del trabajo de un docente con otro. Los exámenes deberían ser el resultado del trabajo realizado por un equipo que persigue los principios de la evaluación, y estándares en elaboración de exámenes.

Los exámenes que sean elaborados deben cumplir las exigencias de la Confiabilidad (reliability), y Validez (validity).

Bibliografía.

- Bermúdez-Jaimes, M., & Mendoza-Páez, A. (2008). La evaluación docente en la pedagogía Montessori: propuesta de un instrumento. *Educación y Educadores*, 11(1), 227-252.
- Careaga, A. (2001). La evaluación como herramienta de transformación de la práctica docente. *Educere*, 5(15), 345-352.
- García, G. (2012). *Investigación comercial*. Madrid: ESIC Editorial.
- González Melo, H., & Ospina Serna, H. (2013). El Saber Pedagógico de los docentes universitarios. *Revista Virtual Universidad Católica del Norte*, 39(1), 95-109.
- Montenegro, I. (2003). *Evaluación de desempeño docente, fundamentos, modelos e instrumentos*. Bogotá: Magisterio.
- Vasco, E. (1990). El saber pedagógico: razón de ser de la pedagogía. En M. Díaz, & J. Muñoz, *Pedagogía discurso y poder*. Bogotá: Corpodic.
- Vidal, D. (2001). *Diseño y elaboración de cuestionarios para la investigación comercial*. Madrid: ESIC Editorial.