

El diseño de exámenes de las clases de español en el ámbito universitario

Resumen del artículo

Trang, Nguyen (2009). "Evaluative Commentary on a Language Test" en *Bulletin of Toyo Gakuen University* 17, pp. 177-193.

Importancia de la evaluación en el aprendizaje y la enseñanza

En la enseñanza de idiomas, la evaluación se puede aplicar para:

- Medir el aprovechamiento
- Evaluar la habilidad
- Tasar los puntos fuertes y débiles de un aprendiente
- Determinar la colocación de un aprendiente en un curso o nivel de estudio.

En un contexto social más amplio, puede servir como:

- Herramienta "disciplinaria" y de "control político".

Objetivo principal de la evaluación

Brindar una medida que sirva de referencia para juzgar la habilidad y competencia de una persona (Hughes 2003:50; Bachman & Palmer 1996: 23).

Requisitos de las evaluaciones válidas y fiables

- Deben demostrar: - que las notas miden con precisión lo que se desea medir
- que las notas mantienen su valor más allá del examen
 - que están relacionadas a competencias lingüísticas particulares
- (Hughes 2003:26; Alderson 1981:55).

En la evaluación de la lengua, resulta crucial que se evalúe a los aprendientes de acuerdo a las destrezas que se desea evaluar y que este proceso se haga a través de textos pertinentes y tipos de pruebas que tengan relación con estas destrezas.

Características de una evaluación apropiada y útil (Según Bachman & Palmer, 1996: 17)

Autenticidad

Fiabilidad

Validez de constructo

Interacción

Viabilidad

Bachman y Palmer indican que para determinar si un determinado ítem o actividad de la evaluación es válido, es

de crucial importancia que se examine la “utilidad” de dicha evaluación y se investigue si cumple con su objetivo de manera que refleje y abarque el aprendizaje de la lengua dentro de un contexto determinado.

Consideran importante que el contenido:

- tenga un propósito claro
- corresponda al uso de la lengua en la “vida real”
- sea diseñado para un grupo específico de aprendientes o participantes.

Análisis de un examen según el modelo de evaluación del lenguaje de Bachman y Palmer (1996)

Propósito del examen:

Examen parcial de segundo año de una escuela técnica de dos años. (4 veces al año)

Objetivo del examen:

Medir el conocimiento de inglés como lengua extranjera durante el curso del semestre académico.

Tipo de evaluación:

De aprovechamiento.

Habilidades evaluadas:

Lectura, escritura, vocabulario

Contexto de aprendizaje:

Objetivo del curso: desarrollo de destrezas para la comprensión de lectura, expresión escrita, expresión oral y comprensión auditiva.

Niveles: 6

Contenido del examen:

Parte del contenido del syllabus de nivel intermedio-básico (bajo)

Refleja los tipos de tareas de escritura y lectura de los libros de texto del curso

Otros métodos de evaluación:

Reportes sobre libros

Pruebas orales

Exposiciones

Participación en clase

Uso de inglés en la clase

Asistencia y esfuerzo

Impacto de la nota final:

Influencia en futuras oportunidades académicas y laborales del estudiante.

Aplicación del examen:

Elaboración: Profesores del programa de inglés

Aplicación: Durante la clase

Duración: 45 minutos

Feedback: Sí, en clases posteriores

Descripción del examen:

Tipo de examen: Escrito

Número de secciones: 8 (A – H) (ver anexo)

Calificación:

Mayor puntaje para la sección de redacción y comprensión de lectura.

Peso del examen:

20% de la nota final.

Características del examen

Fiabilidad

=>Consistencia independientemente del tipo de examen y de las veces que se aplique

- El examen analizado está constituido por un número de secciones cuyas respuestas precisas refuerzan su fiabilidad: desde la sección B hasta la G, se dan respuestas definidas, por lo que independientemente de la persona, lugar y hora a la que se califique, las notas deberán permanecer invariables.
- Los tipos de ítems permiten una calificación fácil:
 - Ítems de respuesta múltiple (secciones F y G),
 - Selección de Sí/ No (sección E)
 - Verdadero / Falso (sección E)

Problema de este tipo de ítems:

- Cabe la posibilidad de que estos ítems no midan el conocimiento de los estudiantes ni su habilidad lingüística, se puede acertar por casualidad.
- Este tipo de ítems implican una “verdad absoluta”, mientras que el proceso de lectura y comprensión de textos en la “vida real” implica la interpretación de múltiples sentidos y formas (Shohamy 2001: 24) Por lo tanto, mientras la facilidad de la calificación puede reforzar la fiabilidad del examen, también puede poner en riesgo su validez.
- Fiabilidad en peligro: En las secciones A y H se tiene que calificar oraciones producidas por los estudiantes, donde la única instrucción es que los párrafos tengan una buena estructura y se empleen oraciones completas. La calificación puede ser subjetiva.
 - Sin embargo, Freedman (1993) y Shohamy (2001) señalan que dichos ejercicios de redacción en los exámenes deben permanecer abiertos y no volverse muy objetivos porque se alejarían de la esencia de la verdadera expresión escrita.
- El contexto de aplicación de los exámenes también puede afectar su fiabilidad:
 - Actitud del profesor ante el fraude (existe una regla general pero su aplicación depende del profesor)
 - Ubicación de los estudiantes en el aula durante el examen
 - Posibilidad de comunicación no verbal entre los estudiantes
 - Permanencia en sus asientos hasta el final del examen
 - Duración del examen
 - Nivel de minuciosidad en las instrucciones

Validez

La validez del constructo concierne al concepto de validez en la evaluación, que también comprende la validez de contenido y la validez de criterio.

Para establecer la validez de un constructo es indispensable identificar la definición del constructo y examinar los ítems o tareas del examen, para poder determinar el grado en el que están relacionados con la interacción y uso

de la lengua en la “vida real”.

En el examen analizado, los constructos son:

- Conocimiento de vocabulario, gramática y uso del idioma inglés a través de la lectura y escritura

El conocimiento se evalúa a través de:

- Ítems de Verdadero / Falso para distinguir entre un hecho real o una opinión (Sección B)
- Ítems para emparejar palabras con sus significados (Sección C)
- Ítems para emparejar palabras con sus sinónimos (Sección D)
- Ítems de Sí / No para distinguir entre verbos, adjetivos y sustantivos (Sección F)
- El vocabulario independientemente se evalúa independientemente de estructuras gramaticales, textos y discursos. →La carencia de un contexto de uso compromete la validez del examen.
- Ítems de respuesta múltiple para evaluar la comprensión de lectura (Sección G) →Reflejan más la identificación de las palabras que la comprensión del texto (i.e.: pregunta 32).

Autenticidad e interacción

Aunque no reemplace eventos de la vida real, lo cual no es su naturaleza ni función, es indispensable que la evaluación incorpore lo aprendido en el aula mientras representa situaciones de la “vida real”.

Spence-Brown (2001): Los contextos de los exámenes son, por naturaleza, artificiales y aunque las pruebas reflejen situaciones de la “vida real”, la manera como los participantes actúan en él se vuelve artificial.

El examen analizado carece de autenticidad e interacción (i.e.: La actividad de la sección F no es auténtica)

Impacto

Frente a los valores y objetivos que representa la evaluación para una sociedad o grupo de personas la idea de objetividad encontrada en ítems dicotómicos: Sí/ No, Verdadero/ Falso, y respuestas correctas o incorrectas que predominan en el examen resulta muy limitada en cuanto a visión, alcance y enfoque (Lynch 2001).

Exámenes de este tipo pueden provocar en los estudiantes una falta de motivación para aprender más de lo que se evalúa.

Aunque se suele considerar que el impacto de la evaluación en las aulas no es muy grande, puede, como en este caso influenciar futuras oportunidades académicas y laborales del estudiante.

Viabilidad

La viabilidad se refiere a la relación entre el desarrollo y la implementación de la evaluación y los recursos disponibles para su realización.

Los recursos pueden ser tiempo, equipo, material, espacio y recursos humanos (Bachman y Palmer 1996)

En este caso, al tratarse de un examen escrito, resulta viable.

CONCLUSIÓN

El examen está constituido en gran parte por elementos discretos y por ello se le critica por adoptar este enfoque en la evaluación.

El énfasis que se da en el empleo de vocabulario sin contexto y de tareas gramaticales para facilitar la calificación sacrificando el uso más profundo e interactivo de la lengua, perjudica la validez del examen.

La composición incluida en el examen sirve para obtener una medida más amplia de la habilidad en la expresión

escrita de los estudiantes que concuerde con el desarrollo de la lengua en el ámbito académico.

Aunque no es posible erradicar completamente las imperfecciones en la evaluación de la lengua (Bachman y Palmer, 1996), se puede reducir los factores que comprometen la validez incorporando tipos de ítems, formatos de respuesta y métodos de evaluación que permiten sacar conclusiones con respecto a las destrezas y habilidades de los individuos (Kobayashi 2002). Resulta esencial que se incluya lo aprendido en el aula y que a la vez se representen situaciones de la “vida real” y usos de la lengua que correspondan al currículo, enseñanza y los objetivos de aprendizaje.

La revisión de este examen empleando la lista de control del modelo de evaluación de la lengua de Bachman y Palmer (1996) ha demostrado que dichos modelos o listas sirven como un marco muy útil para examinar la fiabilidad y utilidad de un examen.

Estos modelos de evaluación de la lengua también demuestran que la validación de la evaluación es un proceso continuo que comienza cuando desde el primer momento de la elaboración de la evaluación y se mantiene como un proceso continuo hasta el final de la misma.

En el núcleo de la evaluación y las pruebas, se encuentra un entendimiento de lo que es la lengua y lo que implica aprenderla y usarla, lo cual se convierte luego en la base para establecer maneras de evaluar la habilidad de la gente.

Una evaluación bien desarrollada sirve para reforzar el ciclo y experiencia de la enseñanza y el aprendizaje, y puede tener un impacto positivo en las funciones educativas, sociales y políticas de la evaluación.

ACTIVIDADES POSTERIORES AL RESUMEN

Se revisaron los distintos métodos y enfoques de enseñanza existentes para confirmar la importancia de tener un modelo de lengua antes de diseñar un examen para que la evaluación de la misma se haga de manera congruente con la concepción de la lengua adoptada en el currículo.

Asimismo, se discutieron los requisitos que debe cumplir un buen examen: fiabilidad, validez, autenticidad, interacción, impacto, viabilidad según Bachman y Palmer (1996) y se mostraron los tipos de ejercicios que se pueden emplear para evaluar la gramática bajo las siguientes clasificaciones: Tareas de respuesta preseleccionada (cerrada), tareas de producción limitada y tareas de producción extensa.

Finalmente, los participantes analizaron las ventajas y puntos débiles de los ejercicios presentados en el artículo tratado, así como las maneras de hacer que los exámenes a nivel universitario tengan mayor fiabilidad.