

**Percepciones de los agentes involucrados en la aplicación de una
prueba de clasificación para alumnos de la extensión del Departamento
de Lenguas de La Universidad Nacional de Colombia**

Claudia Patricia Mojica*

claudiapmojica@yahoo.com

Yeimi Yamile Velásquez Díaz**

yeimivel@yahoo.com

Norma Chavarro Casas***

mnchavac@unal.edu.co

Universidad Nacional de Colombia

Este artículo presenta los resultados de una investigación realizada en la Unidad de Extensión del Departamento de Lenguas Extranjeras de la Universidad Nacional de Colombia. Considerando la problemática generada alrededor de la clasificación de los estudiantes una vez conocidos los resultados, esta investigación tuvo como objetivo principal evaluar la pertinencia del examen de clasificación que los estudiantes que así lo desean toman al ingresar al Programa de Inglés para adultos de la Extensión. El marco teórico que orientó la investigación fue el modelo propuesto por Bachman y Palmer (1996-2000), el cual presenta una perspectiva integral para el diseño y evaluación de exámenes. Los resultados obtenidos indicaron que la utilidad de este examen de clasificación es muy relativa para el contexto.

Palabras clave: evaluación, exámenes, impacto, practicidad

This article shows the results of a research study carried out at the Extension Services at the Foreign Languages Department at Universidad Nacional de Colombia. Taking into account the problems caused by the results of a placement test taken by students at the moment of entering the Adult English Program, research was set out to evaluate its suitability. The theoretical framework that oriented this research was the model proposed by Bachman and Palmer (1996-2000) which presents a very complete perspective for the design and evaluation of tests. The results of this research show that the usefulness of this placement test depends highly on the context.

Key words: Evaluation, tests, impact, usefulness and practicality

* **Claudia Patricia Mojica** es Licenciada en Filología e Idiomas – Inglés de la Universidad Nacional de Colombia. Obtuvo el MA in Applied Linguistics to TEFL de la Universidad Distrital Francisco José de Caldas. Actualmente está vinculada en el colegio Abraham Lincoln como profesora de tiempo completo.

****Yeimi Yamile Velásquez Díaz** es licenciada en Filología e Idiomas – Inglés de la Universidad Nacional de Colombia. Actualmente trabaja como profesora de inglés en los cursos de Extensión del Departamento de Lenguas Extranjeras de la misma universidad y en el colegio Abraham Lincoln.

*****Norma Chavarro Casas** es profesora pensionada del Departamento de Lenguas Extranjeras -Universidad Nacional de Colombia.

Este artículo se recibió en marzo 8, 2009 y fue aceptado para publicación en octubre 28, 2009.

Introducción

La Extensión del Departamento de Lenguas Extranjeras propende por “la continua interacción e integración con la comunidad nacional e internacional”¹ del país y responde a la necesidad sentida en relación con la formación en lenguas extranjeras y áreas afines como la traducción. Es así como la Extensión ofrece cursos para niños, adolescentes y adultos a la vez que exámenes de traducción e interpretación en diversas lenguas.

Antes de iniciar los cursos de inglés, francés y alemán, aquellas personas que llegan con conocimientos de la lengua extranjera tienen la opción de presentar un examen de clasificación que los ubique en el nivel acorde con el grado de dominio de lengua. En reiteradas ocasiones se detectó que un número de estudiantes quedaba en un nivel inferior o superior a su conocimiento teniendo que ser reclasificado después de haberse iniciado el curso. Igualmente se observó que si el estudiante no manifestaba su inconformidad y si el profesor no descubría el caso a tiempo, el estudiante realizaba el curso y, como consecuencia, no aprovechaba al máximo las actividades puesto que le eran o muy fáciles o, por el contrario, muy difíciles, lo cual llegaba a suscitarle sensaciones de fracaso o pérdida de tiempo, a la vez que afectaba, de manera notoria, su autoestima y motivación. Esta situación problemática dio origen a que el Departamento de Lenguas Extranjeras decidiera emprender una investigación que diera luces sobre el estado de correspondencia entre el examen de clasificación y el programa académico propuesto para los cursos de inglés y que, asimismo, pudiera identificar los aspectos que debían modificarse para que el proceso de clasificación tuviera un margen de error más reducido.

Con ese propósito, un grupo de docentes del Departamento inició una investigación cuya primera etapa tuvo lugar en el segundo semestre de 2006. En esta fase se hizo una caracterización del examen de clasificación que corresponde a la serie ‘Channel Your English MM Publications Placement Test’ (ver Anexo 1)² mediante el análisis y contraste de cada pregunta con el programa académico de los diferentes cursos del programa de inglés para adultos. Al finalizar esta primera etapa, se concluyó

¹ Universidad Nacional de Colombia. <http://www.humanas.unal.edu.co/cms.php?id=45> Consultada noviembre 5 2008
² El Anexo 1 contiene una descripción del examen, no el examen en sí.

que en la prueba escrita no existe un orden de complejidad ascendente como tampoco equidad en el número de preguntas que evalúan el conocimiento de los diferentes temas en cada nivel.

Considerando estos resultados y el beneficio que representa para la población estudiantil, el grupo de investigación hizo las siguientes recomendaciones:

- tener en cuenta los objetivos y contenidos de los cursos de extensión,
- adoptar la práctica de evaluación permanente de los diferentes aspectos del currículo presentes en estos exámenes,
- organizar las preguntas de manera secuencial y progresiva, y,
- diseñar el examen atendiendo a los conceptos de validez y confiabilidad que caracterizan a este tipo de pruebas.

Una segunda etapa de la investigación, emprendida por un equipo conformado por dos de los docentes que participaron en la primera etapa y una investigadora principal, tuvo como objetivo continuar la evaluación del examen de clasificación teniendo en cuenta las percepciones de las personas involucradas en el examen, lo cual da a esta investigación un carácter cualitativo, no cuantitativo ni estadístico como tradicionalmente se procede para el análisis de exámenes, y tomando como base el modelo propuesto por Bachman y Palmer (1996-2000). El equipo consideró apropiado utilizar este modelo porque su enfoque presenta una perspectiva integral para el diseño y evaluación de exámenes. A través de seis cualidades (confiabilidad, validez del constructo, autenticidad, interactividad, impacto y practicidad) que los autores ven como complementarias, se puede llegar a determinar el grado de utilidad de un examen. El objetivo de esta segunda etapa de la investigación se concentró en las cualidades de impacto y practicidad. Así, este estudio giró en torno a las siguientes preguntas de investigación:

- ¿Cuáles son las percepciones de los agentes involucrados?
- ¿Cómo está afectando el proceso de aplicación a los participantes?

Se describe, a continuación, el marco teórico, el diseño, los participantes, la recolección de datos, las consideraciones éticas y los resultados obtenidos. Es importante mencionar que por razones de espacio no se mencionan las consideraciones

éticas que se tuvieron en cuenta, ni lo que atañe a la validez y confiabilidad de la investigación.

Marco Teórico

Para abordar las preguntas de investigación, se consideran las percepciones como el proceso de interpretación, valoración y/o significado que estudiantes, profesores y personal administrativo atribuyen a lo relacionado con el examen una vez realizado, y asimismo se toma como marco de referencia para esta segunda etapa del estudio el modelo de evaluación propuesto por Bachman y Palmer (2000). Para conocer si el diseño o un examen es relevante, pertinente y útil para un contexto determinado, las percepciones de las personas involucradas desempeñan un papel primordial. En la valoración de las cualidades del modelo se hace mención especial a éstas.

Con la siguiente figura estos autores, Bachman y Palmer (2000), muestran cómo entienden la idea de utilidad de un examen:

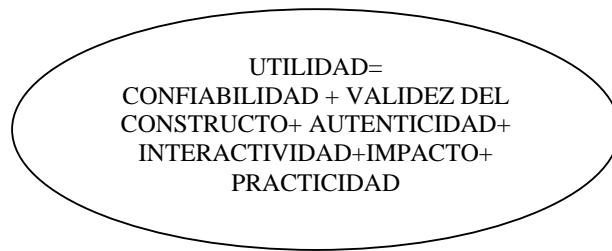


Figura 1. Utilidad de un examen (Bachman y Palmer, 2000)

Es necesario resaltar que los autores hacen énfasis en que las cualidades son complementarias y que es preciso encontrar un balance entre ellas ya sea para el diseño adecuado de un examen o para evaluar su utilidad como en este caso.

A continuación, se hará una breve descripción de las seis características o cualidades consideradas como complementarias por los dos autores³.

³ Es preciso aclarar que el libro consultado está escrito en inglés. Por lo tanto la descripción de las características ha sido interpretada y traducida por las autoras de este artículo.

Confiabilidad: La confiabilidad es frecuentemente definida como la consistencia en los resultados de un examen. Un examen es confiable sí, por ejemplo, es aplicado al mismo grupo de individuos en dos ocasiones diferentes, en dos escenarios diferentes y los resultados son similares. De otro lado, un examen debe medir con precisión aquello para lo cual fue diseñado. En este caso, si se diseña un examen para clasificar a un individuo según su proficiencia en la lengua y éste no lo clasifica en el nivel correspondiente, se dice que el examen no es consistente. Al mismo tiempo es indispensable reconocer que no es posible eliminar por completo las inconsistencias; lo que se puede hacer es minimizar al máximo los posibles efectos de las fuentes de inconsistencia.

Validez del Constructo: Según los autores, ésta corresponde a la relevancia, la conveniencia de la interpretación que se hace de los resultados de los exámenes, es decir, es imperioso demostrar o justificar las interpretaciones de estos resultados. Y para justificar una interpretación de un resultado particular es necesario dar evidencia de que el resultado del examen refleja el área o áreas de la habilidad lingüística que se quiere medir. Para poder dar esta evidencia se debe definir claramente el constructo, es decir, definir las habilidades lingüísticas que se quieren medir. Hay diferentes tipos de evidencia, por ejemplo la relevancia y cobertura del contenido de un examen.

Autenticidad: Este concepto, según Bachman y Palmer (2000), tiene que ver con la relación entre el desempeño en un examen y el uso de la lengua en situaciones diferentes a la del examen. La definen como el grado de correspondencia entre las características de una actividad usada para la prueba y las de una actividad de uso de la lengua. Los autores consideran que esta cualidad es importante por dos razones. La primera porque la autenticidad permite darse cuenta hasta qué punto la interpretación de los resultados puede generalizarse y relacionarse con el uso de la lengua en situaciones externas a las del examen. La segunda razón tiene que ver con el efecto que el examen pueda tener en las percepciones de los individuos y por ende en su desempeño. Una consideración importante para tener en cuenta es que el nivel mínimo aceptable de autenticidad en un examen puede establecerse de dos maneras: a) en términos de las características de las tareas, y b), en términos de las percepciones esperadas por parte de los usuarios del examen (estudiantes y profesores).

Interactividad: Es definida como el grado y la manera como se involucra la persona que toma el examen con las tareas incluidas en éste. Por consiguiente, la interactividad se determina en función tanto de las características de las personas que toman el examen (características personales, habilidad lingüística, conocimiento del mundo y esquema afectivo) como de las características de las tareas del examen. Las características de autenticidad, interactividad y validez del constructo dependen de la definición que se tenga del constructo ‘habilidad lingüística’ en un contexto específico.

Impacto: Otra de las características o cualidades que debe tomarse en cuenta en el momento de determinar la utilidad de un examen tiene que ver con el impacto que el resultado de estos exámenes o la misma experiencia de haberlo tomado genera, no solamente en la persona que toma el examen sino también en otros agentes que están muy involucrados en los procesos de enseñanza y aprendizaje. Se refieren también al efecto ‘backwash o washback’, es decir, el efecto de un examen en los procesos de enseñanza y aprendizaje. Hughes (1989) se refiere a éste como un efecto negativo o positivo; negativo si por ejemplo no hay concordancia entre lo que se enseña y lo que se evalúa; positivo, por ejemplo si hay correspondencia entre las actividades de la clase y las tareas del examen.

Practicidad: Esta última calidad, diferente por naturaleza a las otras cinco, no tiene una conexión directa con la interpretación de los resultados del examen. Sin embargo, este concepto es fundamental dada su relación con aspectos que facilitan la realización del examen y el desempeño satisfactorio de los usuarios. El concepto de practicidad puede llegar a tener efectos sobre el diseño y el tipo de tareas, al igual que el reconsiderar procedimientos que tengan que ver con las otras cualidades que determinan la utilidad del examen. Un examen no es práctico si, por ejemplo, los recursos necesarios para la implementación del examen exceden los disponibles. Es necesario, por lo tanto, considerar y evaluar, de manera rigurosa, todos los recursos (humanos, materiales y de tiempo) a la vez que los costos que se necesitan para la aplicación del examen y la interpretación de los resultados.

Es preciso recordar que, para esta etapa de la investigación, estas dos últimas cualidades fueron el centro de atención.

Diseño de la Investigación

Esta investigación es un estudio de caso de carácter cualitativo evaluativo. Merriam (1998, p. 19), siguiendo a Smith (1978), señala que el estudio de caso cualitativo consiste en una “descripción y análisis intensivos de una unidad única o de un sistema cerrado” y tiene como objetivo “lograr una comprensión a profundidad de la situación y del significado que ésta tiene para las personas involucradas”. A su vez, un estudio de caso evaluativo tiene como finalidad describir, explicar y emitir un juicio de valor ya sea positivo o negativo. La unidad o sistema cerrado de esta investigación es el examen de clasificación.

Participantes

En esta segunda etapa de la investigación participaron: a) los estudiantes que tomaron el examen de clasificación en diciembre de 2006 para iniciar el primer bimestre de 2007, (67); b) los que tomaron el examen el 29 de marzo, para comenzar el segundo bimestre del 2007, (48) para un total de 115 estudiantes; c) docentes del programa que aplicaron en las dos ocasiones la prueba escrita y oral; d) el personal administrativo de la Extensión; y e) una de las investigadoras del equipo, como evaluadora del examen de clasificación en dos ocasiones anteriores a la investigación.

Recolección de Datos

Para dar respuesta a las preguntas de investigación planteadas en esta etapa, se emplearon cuatro procedimientos fundamentales para la recolección de la información: encuestas, entrevistas, notas de campo y retrospección.

En primer lugar, la encuesta a los estudiantes (ver Anexo 2) tuvo con objetivo conocer sus percepciones frente a los resultados del examen y frente al examen tanto escrito como oral. Las encuestas a profesores tuvieron como finalidad conocer sus percepciones en cuanto a la efectividad del examen para clasificar a los estudiantes, los

aspectos que tienen en cuenta en el momento de tomar la decisión para la clasificación y sugerencias que a bien tuvieran dar para el mejoramiento del examen y de los procesos. Respondieron la encuesta cincuenta y tres estudiantes y once profesores.

Las entrevistas fueron de gran ayuda para profundizar en algunos aspectos y como estrategia para mayor validez y confiabilidad de los datos recogidos por las encuestas. Se decidió hacer entrevistas semi-estructuradas por la flexibilidad que permiten. Según Burns (1999, p. 120) este tipo de entrevista tiene la ventaja de facilitar la expresión de la perspectiva que tiene tanto el entrevistado como el entrevistador sobre los objetivos de la investigación y por lo tanto da pie para un balance en la relación investigador-investigado. Estas entrevistas se realizaron a estudiantes que habían quedado mal clasificados, a profesores y a personal administrativo. Se entrevistaron estudiantes, profesores y personal administrativo. En el Anexo 3 se presenta el protocolo de la entrevista a profesores. Las preguntas eran una guía para el entrevistador.

Por otro lado, las notas de campo tomadas por los miembros del equipo investigador durante el proceso de aplicación del examen se emplearon con el fin de observar aspectos de su practicidad y para contrastar la información obtenida a través de los otros dos procedimientos.

Finalmente, uno de los miembros del equipo de investigación hizo retrospección enfocada en las ocasiones en que se le solicitó, como profesora de la Extensión, aplicar el examen. Se optó por este método debido a que por motivos ajenos a la investigación, las entrevistas realizadas a profesores fueron pocas.

Resultados

Después de preparar los datos para su análisis, de realizar el análisis de la información obtenida a través de cada uno de los procedimientos, se cruzaron los datos para compararlos y contrastarlos entre si con el fin de dar mayor confiabilidad al estudio. El análisis tuvo en cuenta las seis cualidades del modelo de Bachman y Palmer, en especial el impacto y la practicidad.

Los resultados presentan: las percepciones de los estudiantes en relación con el resultado del examen en la clasificación y su opinión en cuanto a la parte oral y la escrita; las de los profesores con respecto al conocimiento que tienen de la prueba, su

efectividad para clasificar a los estudiantes y los procedimientos para la clasificación de los estudiantes. A su vez se presentan las sugerencias tanto de profesores como estudiantes.

Percepciones de los Estudiantes frente al Examen de Clasificación

Aunque la mayoría de estudiantes consideran haber quedado bien clasificados por medio del examen, existe un porcentaje significativo: 38% que cree lo contrario. Sin embargo, se detectó, a través de una de las preguntas de la encuesta, que un 12% decía haber quedado bien clasificado aunque consideraba estar más avanzado en algún aspecto de la lengua (gramática y/o fluidez). El grupo de investigación valoró estas respuestas como inconformidades con la efectividad del examen. El siguiente ejemplo tomado de una de las entrevistas así lo confirma:

“Si, porque aunque el comenzar desde los primeros niveles permite recordar bastantes cosas del idioma, podría haberme adaptado a un nivel superior.”

“Porque quedé en un nivel muy básico y ya tengo conocimientos de algunos temas que se han visto.....”

“Porque los temas son bastante básicos y a comparación de algunos de mis compañeros estoy adelantada.”

Es importante observar que en estas respuestas la pertinencia del examen se valora con respecto a los temas tratados en el curso o en relación con el nivel de sus compañeros.

Las razones por las cuales los estudiantes mencionan no haber terminado la parte escrita del examen fueron, en orden descendente en cuanto al número de respuestas: falta de tiempo, extensión del examen, falta de conocimiento y falta de instrucciones claras.

Llamaron la atención estas respuestas:

“debido a la extensión del examen marque la respuesta que no había alcanzado a terminar al pinochazo”⁴

“decidí dejar en blanco aquellos espacios en los que tenía dudas o simplemente no sabía”

⁴Las citas no han sido editadas en ninguno de sus aspectos.

La primera respuesta evidencia lo que usualmente sucede en un examen de opción múltiple. Sin embargo, se infiere que el estudiante no es consciente de la importancia del propósito para el cual toma el examen, lo que sí se demuestra en la segunda respuesta.

En relación con la prueba oral, se consideró relevante caracterizar el ambiente que el entrevistador crea en el momento de la entrevista. Así pues, el mayor porcentaje de respuestas indica que los estudiantes se sintieron en un ambiente cómodo para la entrevista. Entre las cualidades que los estudiantes resaltaron en el docente que aplicó la prueba están: amabilidad, paciencia, dinamismo, comprensión ante los nervios y claridad, tal y como lo evidencia la siguiente respuesta en una de las encuestas:

“A pesar de que fue un examen, en mi caso se formó más como un dialogo muy ameno y mi entrevistadora se mostró muy interesada en facilitar la comunicación”

No obstante, un 34% de los estudiantes, dice no haberse sentido cómodo durante la entrevista. Entre las características negativas sobresalen: el hablar con rapidez, gestos de desaprobación, cansancio, y mal genio que generan nerviosismo e inseguridad. Esto puede observarse en los siguientes ejemplos:

“Estaba nervioso por algún posible error al hablar ya que era evidente algún gesto de desaprobación por parte del evaluador”

“Porque la forma en que me miraba y los gestos hacía que me pusiera muy nervioso”

Tanto las respuestas que apuntan al ambiente positivo así como las que señalan el ambiente negativo son valiosas para esta investigación porque indican que se puede minimizar el impacto negativo fortaleciendo las características positivas en cuanto a la aplicación del examen.

Con respecto a las sugerencias de los estudiantes es preciso mencionar las que tuvieron mayor número de respuestas:

- la necesidad de reducir la extensión del examen
- la importancia de conocer el resultado del examen

Por otra parte, considerando lo consignado en las notas de campo tomadas durante la implementación del examen, se encontró que la falta de información a los estudiantes acerca del tipo examen que van a presentar, el propósito que éste tiene y el

tiempo que dura genera muchas veces un impacto negativo⁵ ya que esto puede causar inseguridad y nerviosismo en el momento del examen y por supuesto, el desempeño del estudiante puede variar, como se evidencia en la siguiente observación de una de las investigadoras:

“De pronto uno de los candidatos salió del salón. Se dirigió a mi y expresó que se sentía muy nervioso, que jamás había presentado un examen igual. Lo tranquilicé, explicándole que el examen no tenía ninguna nota, que era solamente para saber en qué nivel quedaba...”. (Nota de campo de una observación realizada)

Percepciones de los Profesores

Se creería que todos los profesores que aplican el examen lo conocen antes de aplicarlo, sin embargo las encuestas revelaron que no sucede así. Las respuestas dadas por los docentes va desde desconocimiento total del examen a un conocimiento superficial: “acabo de recibir el examen y rápidamente he revisado su contenido” hasta un conocimiento más detallado donde se menciona cada una de las partes de éste. La mayoría de los profesores encuestados había aplicado los exámenes en otra ocasión, solamente el 18% lo aplicaba por primera vez. Por lo tanto, como se observó en las notas de campo hay inseguridad, preocupación y nerviosismo por parte de algunos profesores a la hora de dar instrucciones y liderar el proceso de aplicación del examen.

Con respecto a la efectividad del examen para clasificar a los estudiantes, es importante señalar que de los once profesores que respondieron la encuesta, cuatro piensan que no es efectivo, mientras que tres piensan todo lo contrario, y los cuatro restantes se refieren a que la pauta para la clasificación la da el examen oral.

Otro de los resultados de las encuestas a los profesores tiene que ver con el procedimiento para la calificación y cómo se llega a la decisión de ubicar y clasificar al estudiante en determinado nivel dentro del programa de inglés de los cursos de extensión.

Aquí es necesario recordar que el examen escrito consta de tres partes: en que se evalúa escucha, gramática y lectura. Como este examen se caracteriza por respuesta de selección múltiple hay una sola respuesta para cada pregunta; por lo tanto es relativamente fácil llegar al puntaje. No obstante, no hay unificación en la manera cómo

⁵ Para reducir este impacto negativo, el equipo diseñó un instructivo con información pertinente acerca del examen que incluye objetivo, estructura del examen, tiempo, procedimiento y recomendaciones.

los profesores llegan al total. Algunos escriben el puntaje obtenido en cada parte y otros dan solamente el total de las tres; esto imposibilitaría dar retroalimentación a los estudiantes y dificulta definitivamente una revisión de los exámenes. En este sentido cabe la pregunta: ¿qué influencia tendrá este procedimiento en el momento de tomar la decisión de la clasificación? ¿Dará lo mismo tener en cuenta el resultado total que tener en cuenta el resultado de cada parte?

Como es sabido, la evaluación de la parte oral es más compleja y surgen varias preguntas: ¿Cómo integran los profesores el resultado de esta entrevista con el resultado del examen escrito para llegar a la clasificación del estudiante?

En relación con los aspectos que priman en el momento de clasificar a un estudiante en determinado nivel, los docentes se refirieron a varios aspectos: conocimiento gramatical y competencia comunicativa; las competencias tanto argumentativa como lingüística, manejo de estructuras dentro de una función comunicativa; la integración de las cuatro habilidades básicas, escucha, habla, lectura y escritura; naturalidad, fluidez y manejo de estructuras gramaticales. Todos estos aspectos estarían bien; pero ¿cómo se llevan a la práctica en el momento de la calificación?

Como puede observarse, a través de la investigación surgieron muchas más preguntas que respuestas. Estas preguntas darían pie a continuar investigaciones al respecto.

La dificultad en la toma de decisiones para la clasificación de un estudiante llevó al equipo de investigación a hacer un ejercicio mediante el cual se viera la confiabilidad de los resultados. El ejercicio consistió en lo siguiente. Se hizo una lista con los puntajes reales obtenidos por los estudiantes que presentaron la prueba del 4 de diciembre de 2006 en cada una de las secciones del examen: ‘Escucha’, Parte 1 (P1) y Parte 2 (P2); ‘Gramática’ P1 y P2; ‘Lectura’ P1 y P2; el puntaje parcial y total de cada parte y el gran total. Cada una de las investigadoras, (I1, I2, I3) tomó la lista e hizo la clasificación. Después se compararon los resultados y cada una reflexionó y expuso la manera como había llevado a cabo el proceso de clasificación. Se resalta el hecho de que esta clasificación se hizo con base únicamente en el puntaje del examen escrito.

Este ejercicio ilustra varios puntos:

- No hay criterios claros para que los docentes tomen las decisiones.

- Cada uno puede proceder de distinta manera, lo cual hace que los resultados sean muy subjetivos. Una de las investigadoras hizo una tabla de equivalencias tomando como parámetro el puntaje total, especificando unos valores y asignándole a esos valores un nivel del Programa de la Extensión. Otra de las investigadoras hizo una tabla similar pero tomó en cuenta solamente el puntaje de una de las partes. La otra tuvo en cuenta tanto el puntaje total de cada parte como el gran total. El problema con estas tablas de valores es que todas eran diferentes y finalmente no tenían ninguna relación directa con los contenidos de los cursos del programa de la Extensión.
- Dos de las investigadoras coincidieron en seis casos.
- Las tres investigadoras coinciden en el nivel (2A, 2B, 2C), en cuatro ocasiones.
- Hay mayor confiabilidad si la clasificación la hacen dos o más profesores.
- Al comparar las decisiones tomadas por el equipo de investigación con la decisión de uno de los profesores, quien tuvo en cuenta el resultado del examen oral, no hay concordancia en ningún caso. Es evidente que, al tener en cuenta el examen oral, todos los estudiantes quedaron clasificados en un nivel más bajo. ¿Cuál es el significado de este resultado para la investigación? Surgen muchas preguntas: ¿Está primando el resultado del examen oral para la clasificación? Si esto es así, una persona que tenga conocimientos de gramática, no podría desarrollar la habilidad oral en un curso más avanzado? ¿Qué pasaría si el estudiante es nervioso o si las condiciones del examen oral no fueran las mejores para que el estudiante demostrara todo su conocimiento y habilidad? ¿No sería conveniente que para el examen oral hubiera más de un evaluador, como se hace en los exámenes orales durante el desarrollo de los cursos?

El mismo ejercicio se llevó a cabo en una de las entrevistas a los docentes. Se le plantearon dos casos hipotéticos. Aquí se ejemplifica solamente uno de ellos que equivaldría a la toma de decisión que hace el profesor cuando tiene en cuenta tanto el puntaje del examen escrito como la clasificación del oral. ¿Si un estudiante obtiene un puntaje total de 68 sobre 100 y en el examen oral el evaluador lo clasifica en 1B, ¿en qué nivel lo clasificaría usted?

Al finalizar la encuesta, los docentes dieron sugerencias para mejorar el examen de clasificación.

Entre las respuestas se encuentran, entre otras:

- Disminución de la extensión del examen, lo cual coincide con las sugerencias de los estudiantes,
- renovación del material visual
- inclusión de producción escrita
- preparación de los docentes que aplican el examen
- estandarización de los procesos para la entrevista.

Las sugerencias dadas por los docentes apuntan a minimizar el margen de error en cualidades como la validez del constructo y la practicidad.

Por medio de las observaciones, entrevistas y retrospección se hicieron evidentes ciertos problemas a la hora de la aplicación y calificación del examen, aspectos que influyen negativamente las cualidades de confiabilidad, validez del constructo, impacto y practicidad.

En primera instancia, se encontró que no todos los docentes dan instrucciones en la misma lengua, unos las dan en inglés y otros en español, lo cual redunda en una falta de equidad para los estudiantes, dado que el estudiante que recibió la instrucción en español tiene más ventaja que aquel que la recibió en inglés. Es importante propender por un ambiente de justicia en el cual todos los estudiantes desarrolle el examen bajo las mismas condiciones con el fin de obtener resultados más consistentes.

Este aspecto puede llegar a minimizarse si hay una preparación previa al examen, en donde los profesores puedan llegar a unificar las acciones y/o procedimientos durante la aplicación de la prueba.

Otros aspectos que tienen que ver con el concepto de practicidad, y pueden llegar fácilmente a reducirse en gran medida con una cuidadosa preparación de todos los agentes involucrados en este proceso. Un buen ejemplo de esto tiene que ver la información que debe llegar a toda la población, es decir a estudiantes, profesores, y administrativos, sobre el objetivo, la importancia, las características, el tiempo de duración del examen, el lugar y los materiales. Todos estos aspectos fueron encontrados y reportados, de manera negativa, por varios agentes de la comunidad estudiantil de los cursos de extensión.

Conclusiones

Como se ha podido observar a lo largo de este informe de investigación evaluativa, se puede concluir que las percepciones de los agentes involucrados en la realización de un examen son esenciales para determinar si una prueba es útil para un contexto específico, es decir, si su impacto es positivo, si es confiable, válida, auténtica, interactiva y práctica. En el caso particular de este estudio las percepciones dieron luces sobre diferentes falencias que afectaban cada una de las seis cualidades propuestas en el modelo de Bachman y Palmer (2000) y permitieron constatar que en muchos casos el examen logra clasificar a los estudiantes en un nivel apropiado. Por tal razón, es imposible dar un juicio contundente positiva o negativamente sobre la congruencia entre el examen y el contexto del Programa.

Las conclusiones finales se organizaron de acuerdo con las seis cualidades del modelo empleado para la evaluación del examen de clasificación.

Confiabilidad

- No puede decirse que el examen es totalmente confiable. A pesar de haber un porcentaje alto de estudiantes que se muestra satisfecho con la clasificación (52%), el nivel de no satisfacción resulta ser bastante significativo (38%).
- Los resultados de este examen no pueden considerarse confiables, si se tiene en cuenta que el examen escrito esta diseñado para ser administrado de otra forma. De acuerdo a los autores de este test (MM Placement Test) se debe responder la primera parte y dependiendo del puntaje obtenido (entre 36 y 50), se procede con la segunda parte. Dado que el examen se administra de otra manera (las dos partes se contestan, una después de la otra) la confiabilidad de los resultados no es la misma.

Validez del constructo

- Existe cierta correspondencia entre los constructos (la competencia comunicativa y las habilidades) que pretende desarrollar el programa y (la competencia y las habilidades) que mide el examen. El examen escrito se refiere a los niveles del Marco Común Europeo y los libros de texto utilizados por el

programa están diseñados con miras a unificar estos niveles. El examen oral se centra en los niveles de los cursos de Extensión. La no congruencia se manifiesta precisamente en este hecho.

- Los niveles de las bandas del examen MM Placement Test no corresponden a la distribución de los niveles del currículo de los cursos del programa. El Programa consta de doce niveles y el Marco de Referencia Común Europeo de seis: A1, A2, B1, B2, C1 y C2. No hay claridad sobre la correspondencia de niveles entre uno y otro, lo cual afecta tanto la validez como la confiabilidad.

Autenticidad

Aunque hay dificultad en orientar el Programa hacia un uso específico de la lengua por la heterogeneidad de la población, diseñar tareas auténticas no es tan fácil. Si se piensa en el objetivo de las tareas del examen, puede decirse que son relativamente auténticas en el sentido de semejar situaciones de uso de la lengua (entrevista, situaciones de uso de la lengua en el aula, etc.) y relativamente útiles. Sin embargo una de las razones para evaluar esta calidad, según los autores, se relaciona con el efecto que el examen pueda tener en las percepciones esperadas de los usuarios, y por ende en su desempeño. El hecho de que los estudiantes que quedan mal clasificados, se sientan inconformes con su desempeño puede verse como un indicio de falta de autenticidad del examen para la población.

- La diferencia entre los procedimientos y criterios de evaluación empleados por los profesores de la Extensión en sus cursos de inglés y los utilizados para la evaluación en el examen de clasificación son también señales de una falla en la autenticidad de la prueba.

Interactividad

- Se observó que los individuos que toman el examen se involucran en contestarlo y que las tareas en general comprometen al estudiante con su conocimiento de la lengua. Sin embargo, se pudo ver que tanto docentes como estudiantes sugerían replantear la extensión del examen. El cansancio producido por la longitud y duración del examen, al igual que por la hora en que se realiza, hace que el desempeño de estudiantes y profesores no sea el mejor, afectando así los resultados del examen.

Impacto

- Se puede concluir que los efectos de la experiencia de tomar examen, de las decisiones tomadas a partir de los resultados obtenidos, del proceso de implementación del examen y del examen mismo generan un impacto negativo en los estudiantes en cuanto a la no satisfacción en su desempeño en el examen y en la inconformidad con la clasificación; o positivo en lo referente a la efectividad del examen.

Practicidad

- Como se mencionó, falta unificar criterios para los procesos de implementación de la prueba que propendan por un ambiente de justicia y equidad en el desarrollo del examen.
- La no preparación de los docentes para la implementación y evaluación tanto de la prueba escrita como de la oral dificulta la toma de decisiones para la clasificación.
- La falta de información clara y precisa a estudiantes, profesores, y administrativos sobre aspectos que tienen que ver con el concepto de practicidad: objetivos, características e importancia del examen; tiempo de duración, el lugar y materiales afecta la disposición, el desempeño y las decisiones tomadas no garantiza la confiabilidad del examen y por ende sus resultados y las decisiones que se toman a partir de éstos.

En resumen se puede decir que es preciso revisar todas las conclusiones con el fin de minimizar el impacto negativo que el examen genera en los diferentes miembros de la comunidad.

Se puede concluir que la utilidad de este examen para el contexto del programa de Extensión de la Universidad Nacional de Colombia es muy relativa. Es imposible llegar a decir que un examen, cualquiera que sea su propósito cumpla con las seis cualidades al 100%. Se deben recordar las palabras de los autores:

“In our experience, many people believe, as we did, that there is an ideal of what a “good” language test is, and want to know how to create tests on this ideal model for their own testing needs. Our answer is that there is no such thing as a “good” or

“bad” test in *the abstract*, and that there is no such thing as the one “best” test even for a specific situation” (Bachman, 2000, p. 6).

Sin embargo, cada institución debe procurar que todos los elementos que componen el currículo sean cuidadosamente planeados, llevados a la práctica y evaluados periódicamente con el fin de garantizar un sistema educativo de alta calidad.

Referencias

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2000). *Language testing in practice*. Oxford: Oxford University Press.
- Burns, A. (1999) *Collaborative action research for English language teachers*. Cambridge: Cambridge University Press.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education. Revised and expanded from case study research in education*. San Francisco CA: Jossey-Bass Inc. Publishers.
- Smith, L. M. (1978). An evolving logic of participant observation, educational ethnography and other case studies. In Shulman (Ed.), *Review of Research in Education*. Itasca, III.: Peacock, 1978. In Merriam, S. B. (1998). *Qualitative research and case study applications in education. Revised and expanded from Case study research in education*. San Francisco CA. Jossey-Bass Inc. Publishers.

Anexo 1: Descripción del examen

El examen de clasificación vigente en el momento de realización de esta investigación es el *MM Publications* 2004, editorial de la serie *Channel*, cuyos libros de texto se habían implementado en todos los cursos del Programa de Inglés para Adultos y que desde la adopción de la serie *Upstream* de *Express Publishing* (2002) y actualmente de la serie *Total English de Cambridge* no se están utilizando.

1. ESTRUCTURA DEL EXAMEN

El examen de clasificación consta de una prueba escrita y una prueba oral.

a. Prueba escrita

Ésta consta de dos partes. Cada parte incluye un total de cincuenta preguntas de selección múltiple con tres opciones de respuesta y está compuesta por tres secciones, identificadas como A, B y C, las cuales evalúan habilidades específicas.

La siguiente tabla muestra la forma como están distribuidas las preguntas en cada parte de la prueba escrita:

SECCIÓN	HABILIDAD EVALUADA	PREGUNTAS
A	Comprensión auditiva	1 a 10
B	Vocabulario y gramática	11 a 30
	Comunicación	31 a 40
C	Comprensión de lectura	41 a 50

Si bien las dos partes de la prueba escrita mantienen la misma estructura y evalúan las mismas habilidades, en la parte II el nivel de complejidad de las preguntas es mayor y se requiere un conocimiento más avanzado del uso de la lengua por parte del evaluado.

b. Prueba oral

Existen dos formatos de entrevista que denominaremos A y B.

El formato A corresponde a un examen llamado “Spectrum Placement Test” y esta conformado por treinta y tres preguntas organizadas ascendente de acuerdo con estructuras gramaticales utilizadas. Los niveles de clasificación van desde el 1B hasta 3A. Este formato incluye un “checklist” de estructuras gramaticales el cual se usa para llevar una estadística de los conocimientos y debilidades del estudiante para tener una idea más clara del nivel correspondiente al momento de la clasificación; incluye 5 tarjetas con dibujos y una lista de treinta y tres preguntas para utilizar esas tarjetas. Adicionalmente incluye un procedimiento para la administración de la entrevista con recomendaciones para tener en cuenta en el momento de realizar la prueba.

El formato B consta de dieciocho preguntas, organizadas en una escala ascendente de acuerdo con el nivel de complejidad que se requiere en las respuestas y con base en la cual se debe clasificar al aspirante según el número correcto de preguntas que haya respondido de acuerdo con los parámetros lingüísticos especificados. Los niveles de clasificación van desde 1B hasta 4A. No aparecen preguntas para clasificar a un aspirante al nivel 3B porque es un tipo de curso de consolidación en el que se hace una revisión de los temas vistos en los cursos anteriores.

Anexo 2: Encuesta a estudiantes clasificados a través del examen

**UNIVERSIDAD NACIONAL DE COLOMBIA
DEPARTAMENTO DE LENGUAS EXTRANJERAS
EXTENSIÓN**

Encuesta a estudiantes clasificados a través del examen

Apreciado estudiante:

El Departamento de Lenguas Extranjeras de la Universidad Nacional de Colombia está llevando a cabo una investigación cuyo objetivo es analizar la pertinencia del examen de clasificación que se aplica en los cursos de Extensión a la luz de la orientación académica del programa.

Para este fin solicitamos a usted de la manera más cordial responder esta encuesta. La información recolectada será utilizada con fines investigativos y será de carácter estrictamente confidencial.

I. INFORMACIÓN PERSONAL

- Nombre: _____
- Nivel que cursa actualmente _____
- Ha realizado estudios de inglés anteriormente? Si _____ No _____
- Si su respuesta es afirmativa, en donde? _____
Por cuánto tiempo? _____

II. PERCEPCIÓN DEL RESULTADO DEL EXAMEN DE CLASIFICACIÓN

- ¿Considera que cuando usted fue clasificado por medio del examen, el nivel en el cual fue ubicado es/era apropiado de acuerdo con sus conocimientos del idioma inglés?

Sí _____ No _____ ¿Por qué?

-
- ¿Si piensa que el nivel no es/era el apropiado para usted, ¿cree que debió haber sido clasificado en un curso superior, o inferior? ¿Por qué?
-

III. PERCEPCIÓN DEL EXAMEN ESCRITO

- ¿Contestó todo el examen? Sí _____ No _____
- Si su respuesta fue negativa, marque con una X la opción u opciones que usted considere que más se ajusten a su caso.

- _____ Falta de tiempo
- _____ Otro(s) ¿Cuál(es)? _____
- _____ Falta de conocimiento del idioma
- _____ Falta de claridad en las instrucciones
- _____ Presentación del examen (tamaño de letra, formato del examen)
- _____ Extensión del examen

- ¿Respecto a la sección de comprensión auditiva, piensa usted que el grado de dificultad de la parte dos difería de la parte uno?

Sí _____ No _____ ¿Por qué?

IV. PERCEPCIÓN DEL EXAMEN ORAL

- ¿Cómo se sentía antes de comenzar la prueba? Marque con una X la opción u opciones adecuadas a su caso particular:

- _____ Tranquilo(a)
- _____ Inseguro(a)
- _____ Otro(s) ¿Cuál (es)? _____
- _____ Confiado(a)
- _____ Nervioso(a)

Preocupado(a)
 Ansioso(a)

¿Por qué? _____

- ¿La forma en que el entrevistador llevó a cabo el examen contribuyó a que usted se sintiera cómodo?
Si _____ No _____

¿Por qué? _____

- ¿Fueron motivantes las ayudas visuales (dibujos) usadas para llevar a cabo el examen?
Si _____ No _____ ¿Por qué?

V. SUGERENCIAS

Gracias por su colaboración.
Cordialmente,
GRUPO DE INVESTIGACIÓN

Anexo 3: Protocolo entrevista a profesores

UNIVERSIDAD NACIONAL DE COLOMBIA
DEPARTAMENTO DE LENGUAS EXTRANJERAS
EXTENSIÓN

Protocolo entrevista a profesores

Introducción

El Departamento de Lenguas Extranjeras está llevando a cabo una investigación cuyo objetivo es analizar la *pertinencia* del examen de clasificación que se aplica en los cursos de Extensión *a la luz de la orientación académica del programa*.

Para este fin queremos hacerle algunas preguntas. La información dada será estrictamente confidencial y su uso exclusivamente para los fines del proyecto.

Preguntas

- ¿Cómo fue el proceso de asignación como profesor para la aplicación del examen de clasificación?
- ¿Qué información previa al examen le dieron? ¿Quién se la dio?
- ¿Qué función desempeñó usted en el examen? ¿Quién asignó dichas funciones?
- Si un profesor evalúa la parte escrita del examen y otro la parte oral, ¿cómo se toma la decisión del nivel en que se clasifica ese estudiante?
- Supongamos que un estudiante obtuvo 41 puntos sobre 50 en la primera parte y 23 sobre 50 en la segunda, ¿cómo lo clasificaría sin tener en cuenta el examen oral?
- Vamos a suponer que un estudiante tiene un puntaje de 68 sobre 100 en la parte escrita, y el entrevistador lo clasificó en el nivel 1B, ¿cómo hace para determinar en qué nivel queda el estudiante?
- En su opinión, ¿qué determina la clasificación de un estudiante: la parte escrita o la parte oral?
- Para la entrevista, ¿cuál de los dos formatos utiliza: A o B?
- ¿Ha realizado alguna vez un examen de reclasificación? ¿Cómo ha sido el proceso?

Número 3



Diciembre de 2009