

UNIVERSIDAD IBEROAMERICANA

Estudios con Reconocimiento de Validez Oficial
por Decreto Presidencial del 3 de abril de 1981



“ANÁLISIS DEL PROCESO DE COLOCACIÓN PARA DETERMINAR EL
NIVEL DE INGLÉS DE LOS ASPIRANTES A LOS CURSOS DEL
CENTRO DE IDIOMAS DE LA UNIVERSIDAD IBEROAMERICANA,
CIUDAD DE MÉXICO”

TESIS

Que para obtener el grado de

**MAESTRA EN
INVESTIGACIÓN Y DESARROLLO DE LA EDUCACIÓN**

Presenta

Alejandra Aguirre Castañares

Director: Mtro. Francisco Alvarado García

Lectores:

Dr. Javier Loredó Enriquez
Mtro. Jorge Martínez Sánchez

RESUMEN

El presente estudio analiza la efectividad, validez y confiabilidad del proceso de clasificación que se lleva a cabo en el Centro de Idiomas de la Universidad Iberoamericana, Ciudad de México, para establecer el nivel del idioma inglés de los aspirantes a los cursos impartidos en la Dirección de Educación Continua.

El estudio se realizó, tanto desde una perspectiva cualitativa, como desde una cuantitativa.

El aspecto cualitativo de la investigación permitió conocer el proceso y su resultado, y reportó evidencia de su validez desde el punto de vista de los 249 evaluados en el periodo de Primavera de 2007. El resultado de los instrumentos utilizados mostró que todos los sujetos que respondieron consideraron haber sido ubicados en el nivel correcto. El seguimiento que se le dio a los estudiantes que se inscribieron a los cursos indicó que sólo se requirió el cambio de nivel de una estudiante.

La sección cuantitativa se enfocó fundamentalmente al análisis estadístico de la prueba escrita aplicada a 558 candidatos que la respondieron en los periodos de Primavera, Verano y Otoño de 2006, así como Primavera de 2007. El resultado arrojó un grado de confiabilidad muy alto: 0.98. El índice de dificultad fue de 0.41, (lo que lo hace un examen relativamente difícil), y los de discriminación, tanto entre los sujetos, como el de la relación entre los reactivos y el nivel de dificultad, fueron de 0.65 y 0.48, respectivamente, es decir, estuvieron en un rango de *excelente*.

Se hacen recomendaciones para mejorar la evidencia de validez y confiabilidad de la sección oral que forma parte del examen de colocación.

INDICE GENERAL

RESUMEN	2
ÍNDICE GENERAL	3
LISTA DE TABLAS	6
<u>INTRODUCCIÓN</u>	8
▪ Antecedentes	8
▪ Objetivo	9
▪ Metodología	9
▪ Estructura del documento	9
<u>I. PLANTEAMIENTO DEL PROBLEMA</u>	12
▪ La clasificación estudiantil en las clases de lenguas	12
- Sistemas, mecanismos e instrumentos de clasificación	13
▪ Problemática de la clasificación estudiantil	15
▪ Preguntas de investigación	16
▪ Objetivo	17
▪ Justificación	17
<u>II. MARCO REFERENCIAL</u>	23
▪ Aspectos lingüísticos	23
- El lenguaje	23
- Aprendizaje vs. Adquisición, una base para la elaboración de exámenes	25
- La competencia y el dominio lingüístico	29
▪ Evaluación educativa	31
- La evaluación del proceso enseñanza-aprendizaje	34
- Medición del aprendizaje	35
- Confiabilidad	36

- Validez	38
- Relación entre confiabilidad y validez	46
▪ Evaluación del aprendizaje de lenguas	47
- Objetivos de los exámenes de lengua:	49
De progreso	49
De logro	49
De dominio	49
De diagnóstico	50
De colocación	50
- Tipos de exámenes:	51
Subjetivos vs. Objetivos	51
Directos vs. Indirectos	52
De Puntos discretos vs. Evaluación Integradora	52
Referentes a un Criterio vs. Referentes a la Norma	53
De velocidad vs. de poder	53
▪ Consideraciones adicionales	54
- El examen y el proceso de colocación	54
- Los exámenes de opción múltiple	61
- La evaluación oral y su problemática	64
<u>III. MARCO CONTEXTUAL</u>	70
▪ Descripción del Centro de Idiomas de la Dirección de Educación	
Continua de la UIA, del examen escrito y de la entrevista oral.	70
<u>IV. METODOLOGÍA.</u>	77
▪ Tipo de investigación (enfoque)	77

▪ Diseño	77
▪ Unidad de análisis (muestra)	78
▪ Instrumentos de recolección de datos	78
▪ Fases de la investigación	80
- Inicios	80
- Desarrollo	81
- El proceso de clasificación de Primavera 2007	83
<u>V. ANÁLISIS DE RESULTADOS</u>	89
▪ Reporte cualitativo de resultados del proceso de clasificación	89
▪ Reporte cuantitativo de resultados del instrumento escrito	102
- Análisis de reactivos	111
<u>VI. DISCUSIÓN</u>	127
▪ Implicaciones	127
▪ Conclusiones	128
▪ Recomendaciones	134
▪ Comentario final	138
<u>BIBLIOGRAFÍA</u>	139
<u>APÉNDICES</u>	148
▪ Hoja de información del aspirante.	149
▪ Cuestionario de salida para aspirantes.	151
▪ Cuestionario para profesores.	152
▪ Cuestionario para alumnos.	153
▪ Matriz Conducta / Contenido.	154
▪ Tabla de especificaciones.	161
▪ Análisis de reactivos con <i>ITEMAN 2.0</i> .	167

LISTA DE TABLAS

1. Escala de dominio lingüístico MCERL	Pág. 30
2. Escala de valoración analítica	68
3. Panorama de cursos actuales impartidos por la UIA	71
4. Cuestionario para profesores (Piloto, otoño 2006)	82
5. Cuestionario para alumnos (Piloto, otoño 2006)	82
6. Candidatos que se inscribieron	89
7. Lugar de inscripción	90
8. Curso a inscribir	90
9. Curso a inscribir / No. de inscritos: Crosstabulation	91
10. Curso a inscribir / Curso en el que participa: Crosstabulation	92
11. Cuestionario 1. a) ¿El proceso te ubicó en el nivel adecuado?	93
12. Cuestionario 1. c) Si repitieras el proceso mañana sin estudiar ¿obtendrías = result?	94
13. Razones por las cuales, si repitieran el proceso, no obtendrían el mismo resultado	94
14. Cuestionario 1. e) De acuerdo con el resultado ¿es tu intención inscribirte al curso?	95
15. Razones para no inscribirse al curso	95
16. Nivel de ubicación	96
17. Nivel de ubicación / Nivel de autoevaluación con base en MECRL: Crosstabulation	97
18. Cuestionario 2. ¿Consideras que el nivel es correcto?	98
19. Si no es así ¿consideras tu nivel más alto o más bajo?	99
20. Comentarios de alumnos que consideraron estar mal clasificados	99
21. Profesores que detectaron alumnos con diferente nivel al impartido	100
22. Estadística descriptiva por nivel	104
23. Índice de dificultad (Lafourcade, 1973)	107
24. Índice de discriminación (Ebel y Frisbie, 1986)	108

25. Estadística descriptiva prueba completa	109
26. Análisis de reactivos fuera de los criterios de dificultad y discriminación	112
27. Análisis de reactivos: Elemental 11	114
28. Análisis de reactivos: Intermedio 23	114
29. Análisis de reactivos: Intermedio-Alto 8	115
30. Análisis de reactivos: Intermedio-Alto 11	115
31. Análisis de reactivos: Intermedio-Alto 20	116
32. Análisis de reactivos: Intermedio-Alto 23	116
33. Análisis de reactivos: Intermedio-Alto 24	117
34. Análisis de reactivos: Avanzado 1	117
35. Análisis de reactivos: Avanzado 3	118
36. Análisis de reactivos: Avanzado 5	118
37. Análisis de reactivos: Avanzado 8	119
38. Análisis de reactivos: Avanzado 9	119
39. Análisis de reactivos: Avanzado 15	120
40. Análisis de reactivos: Avanzado 16	120
41. Análisis de reactivos: Avanzado 17	121
42. Análisis de reactivos: Avanzado 21	121
43. Análisis de reactivos: Avanzado 22	122
44. Análisis de reactivos: Avanzado 23	122
45. Análisis de reactivos: Avanzado 24	124
46. Análisis de reactivos: Avanzado 25	123
47. Análisis de reactivos: Avanzado 27	124
48. Análisis de reactivos: Avanzado 29	124
49. Análisis de reactivos: Avanzado 30	125
50. Porcentaje de distribución del índice de dificultad (real e ideal) de la prueba	126

INTRODUCCIÓN

▪ Antecedentes

El presente trabajo está relacionado con un elemento esencial en la educación: la evaluación. Aún cuando la evaluación siempre ha existido, es desde hace relativamente pocos años que ha tomado auge y se le ha otorgado la importancia debida (Rueda y Nieto, 1994). Hoy en día existe mayor investigación al respecto, así como diversos exámenes a nivel internacional con una base científica que permiten determinar el nivel de aprovechamiento y aprendizaje de los estudiantes en las distintas áreas evaluadas (Cfr. CENEVAL, INEE, PISA, TIMM).

Evaluar quiere decir recolectar la información necesaria para hacer un juicio sobre la calidad o valor de algo con el fin de tomar decisiones para mejorar. En el caso de la evaluación educativa los propósitos pueden ser muy diversos, pues en este ámbito podemos interesarnos en evaluar el aprendizaje de los estudiantes, un programa, a los docentes, un centro de enseñanza, etc. De una u otra forma y de acuerdo con La Junta de Andalucía (1996), la principal función de la evaluación es “la de apoyo y orientación para la mejora de las acciones educativas, la de ayudar a cada alumno y a cada alumna a superar sus dificultades y a aprender mejor.”

Por lo que se refiere a la evaluación del inglés como segunda lengua (L2) o como lengua extranjera (LE), el uso general de los exámenes es determinar fortalezas y debilidades en las habilidades aprendidas (en LE) y/o adquiridas (en L2) por los estudiantes.

Al igual que en cualquier tipo de evaluación, los distintos exámenes utilizados en la enseñanza de una lengua extranjera deben poder proporcionarnos medidas válidas y confiables enfocadas a su propósito específico. Tal es el caso de los exámenes de colocación como el que revisaremos aquí, cuyo objetivo es identificar el desempeño de un candidato para ubicarlo en el nivel adecuado de instrucción dentro de un programa de estudios.

Sin embargo, no podemos limitarnos a analizar únicamente la prueba escrita, puesto que no es el único elemento que se utiliza para determinar el nivel de idioma de una persona. Por lo tanto, revisamos el proceso de colocación completo. Esto incluye las circunstancias alrededor del proceso, la forma de aplicación de los profesores, su interrelación con los estudiantes, y cómo se realiza la entrevista. Asimismo, estadísticamente analizamos el nivel de confiabilidad, el rango de discriminación entre los sujetos por nivel, y el índice de dificultad de la prueba escrita.

▪ **Objetivo**

El propósito de llevar a cabo esta investigación es determinar si el proceso y los instrumentos que se utilizan en el Centro de Idiomas de la Universidad Iberoamericana para establecer el nivel de un candidato y el curso al que debe inscribirse, es confiable y si presenta evidencia de validez.

▪ **Metodología**

El presente es un estudio mixto-dominante porque utiliza tanto el método cualitativo, como cuantitativo para el análisis. A la vez que se realizan observaciones para describir el proceso y se recaban las opiniones de los usuarios, se recurre a la estadística para el análisis de la prueba escrita para evaluar su grado de precisión.

▪ **Estructura del documento**

El presente documento está organizado en 6 capítulos. En el primero se establece la problemática de la clasificación estudiantil y se exponen las preguntas de investigación, el propósito del estudio y la justificación del mismo.

En el segundo capítulo, “Marco Referencial”, se hace una revisión de la literatura y se divide en cuatro secciones: Aspectos lingüísticos, Evaluación educativa, Evaluación del aprendizaje de lenguas, y Consideraciones adicionales.

En la primera se hace mención de lo que es el lenguaje, se analiza la diferencia entre *aprendizaje* y *adquisición*, conceptos relevantes para la elaboración de exámenes, y la competencia y el dominio lingüístico. En la segunda se revisa el significado de *evaluar*, así como los términos más importantes alrededor de ello, como son la medición del aprendizaje, y la importancia de la confiabilidad y la validez, así como de la relación entre ellas. La tercera parte trata de manera específica sobre los objetivos de los exámenes de lengua y los tipos de exámenes de idiomas. Por último se profundiza en el examen y el proceso de colocación, el funcionamiento de los exámenes de opción múltiple, y la problemática de la evaluación oral.

El “Marco Contextual” es el tema del capítulo tres, y en él se describe el Centro de Idiomas, el origen del examen escrito, la entrevista oral, y cómo funciona el proceso de clasificación en general.

En el capítulo “Metodología”, el número cuatro, se establece el tipo de investigación, el diseño y la muestra, los instrumentos utilizados para la obtención de datos, así como las fases del estudio. En esta sección se describe también el proceso de clasificación como ocurrió en el periodo de Primavera de 2007.

El quinto capítulo da cuenta de los resultados cualitativos, es decir, los relacionados con el proceso de clasificación, así como de los cuantitativos, esto es, la confiabilidad de la prueba escrita, su nivel de dificultad y grado de discriminación. Asimismo se realiza un análisis de aquellos reactivos que hayan obtenido índices bajos en estos aspectos.

En el capítulo final se exponen las conclusiones del estudio y se hacen algunas recomendaciones consideradas pertinentes tanto para mantener la calidad del proceso, como para mejorar algunos aspectos del mismo.

I. PLANTEAMIENTO DEL PROBLEMA

▪ La clasificación estudiantil en las clases de lengua

En el caso de la enseñanza de lenguas y de acuerdo con el Marco Común Europeo de Referencia para las Lenguas (MCERL, 2002) la *evaluación* se refiere en sentido concreto a la valoración del grado de dominio lingüístico que posee un usuario de algún idioma específico. El *Marco* establece la existencia de distintos tipos de evaluación como son de aprovechamiento vs. dominio; formativa vs. sumativa; directa vs. indirecta; de la actuación vs. de conocimientos; subjetiva vs. objetiva; con referencia a la norma o a un criterio; entre otros.

Uno de los conceptos que nos interesa especialmente para este trabajo es la *evaluación del dominio*, que se refiere a “lo que alguien sabe o es capaz de hacer en cuanto a la aplicación en el mundo real de lo que ha aprendido” (MCERL, 2002). Es decir cuando se evalúa el nivel de dominio de un usuario se puede determinar dónde está con respecto a la competencia lingüística establecida ya sea en un programa de estudios, o en relación con un Marco de Referencia como el que citamos aquí.

De acuerdo con el DRAE (2000) se entiende por *clasificación* el ordenar o disponer por clases. En las escuelas de idiomas la clasificación permite analizar las habilidades de los estudiantes para comunicarse en una lengua extranjera y así colocarlos en los niveles de instrucción establecidos en un programa específico (*Foreign Language Framework for California Public Schools*, 2003).

Un programa de calidad en el que los estudiantes se mueven en forma secuencial en cierta lengua extranjera, deberá contar con un proceso de clasificación que permita determinar cuál es el mejor lugar para que el estudiante comience su instrucción.

El proceso de clasificación de los alumnos que ingresan a un programa de estudios de lengua tiene como objetivo evaluar los niveles de habilidad lingüística con el fin de colocarlos en el curso o clase apropiada.

La clasificación de los alumnos por grado o calidad está relacionado con las evaluaciones *con referencia a la norma*. Esto supone que la valoración y la posición de cada persona se realiza con respecto a la de sus compañeros. Un uso común de la evaluación *referida a la norma* ocurre en los exámenes de colocación para formar grupos de los distintos niveles, ya que la *referencia a la norma* se entiende en relación con *la clase* (alguien es el número 2, el 7, etc.) o en relación con el número de alumnos que realiza la prueba. Cuando la referencia es a un *criterio* se parte de un cuadro en cuyo eje vertical se sitúa el grado de dominio lingüístico (como en una línea continua) y en cuyo eje horizontal se recoge la serie de ámbitos relevantes, de manera que los resultados individuales de una prueba puedan situarse en relación con el mapa total de criterios. (MCERL, 2000)

- Sistemas, mecanismos e instrumentos de clasificación

Hemos dicho que para ingresar a un programa de estudios de una lengua en específico se requiere información sobre el dominio gramatical del estudiante así como del uso integrado que hace de la lengua, especialmente en la habilidad oral. Es por esto que por lo general los sistemas de clasificación para el ingreso a programas de idiomas están constituidos al menos por dos elementos: una prueba sobre papel y una entrevista en la lengua meta (LM). Por lo general se considera que los exámenes escritos pueden contener cierto grado de error, y por esta razón se realizan las entrevistas orales para poder asegurar al máximo una clasificación correcta (Fulcher, 1997). Con ambas bases se pueden determinar de manera más precisa las habilidades lingüísticas del estudiante (*Foreign Language Framework for California Public Schools*, 2003).

Además de los típicos exámenes sobre papel hoy en día y gracias a la tecnología se han desarrollado también exámenes de clasificación que se realizan a través de la *Web*, como es el caso de la prueba del Centro de Lenguas de la Universidad de Stanford. Este tipo de exámenes de colocación en línea tiene la ventaja de evaluar al interesado mucho antes de su llegada al campus o instituto de lenguas, y puede ser bastante práctico en cuanto ahorro de tiempo, ya que aplicar la prueba gramatical en línea permite dedicar tiempo a las entrevistas en LM que se realizan frente a frente y que, como ya hemos dicho, pueden ser determinantes en la clasificación. (Bernhardt, et al 2004).

Asimismo, podemos considerar de manera general dos tipos de exámenes de clasificación. Los estandarizados, es decir, pruebas elaboradas con base en el orden de adquisición que establece la investigación en Adquisición de Segunda Lengua¹. Desafortunadamente pueden presentar dos desventajas: una es que el precio puede ser casi prohibitivo, por lo que no todas las instituciones tienen la posibilidad de costearlo (Hinkelman y Grose, 2004). Tal es el caso del *Quick Placement Test* de Oxford University Press cuyo propósito es proporcionar a los profesores un método confiable y rápido de determinar el nivel de inglés de los estudiantes. La segunda desventaja de este tipo de exámenes es que, de acuerdo con Green y Weir (2004), no todos los programas de estudio coinciden necesariamente con el orden establecido por la investigación en Adquisición de Segunda Lengua, lo que puede tener como consecuencia que la clasificación de los estudiantes a través de los exámenes estandarizados, no vaya de acuerdo con el programa utilizado por la institución.

El segundo tipo de exámenes es el de los llamados *hechos en casa*, es decir, creados directamente en los centros de lenguas y que tienden a basarse en sus propios programas de estudio para asegurar una mejor clasificación, pero que en ocasiones pueden enfrentar problemas

¹ SLA por sus siglas en inglés: Second Language Acquisition research.

de confiabilidad y validez ya que no siempre son elaborados por expertos o validados estadísticamente, como el caso del examen de VELI (Vancouver English Language Institute) del que da cuenta Murray, 1999.

▪ **Problemática de la clasificación estudiantil**

En torno a los exámenes de colocación pueden presentarse diversos problemas. Uno de ellos puede surgir si se utiliza un instrumento diseñado con otro propósito. Por ejemplo un examen como el *Test of English as a Foreign Language* (TOEFL) no podría emplearse para determinar el nivel al que un estudiante debe ingresar en un programa de estudios, aún cuando su objetivo sea determinar el dominio de la lengua del alumno. La validez de un examen es imprescindible y esto implica que el instrumento mida lo que pretende medir, y nada más, es decir, que no se use para otros fines. Asimismo, los exámenes deben ser confiables, esto es, deben mostrar estabilidad y consistencia en las puntuaciones cada vez que se utilizan. El problema más común con los exámenes de colocación es la falta de precisión para ubicar a los interesados en el nivel correcto y esto puede deberse a distintos factores como son: un examen mal diseñado en cuanto a instrucciones, reactivos, y elementos lingüísticos a evaluar; la falta de capacitación de los examinadores por ejemplo en cuanto a los criterios para realizar una entrevista LM; y la falta de una adecuada logística por parte de los administradores (costo, horario, lugar, duración, etc).

Las consecuencias que una clasificación inadecuada puede traer consigo incluyen problemas a nivel del estudiante, a nivel administrativo y del programa. En el caso de los estudiantes, cuando no se percatan del problema pueden sentirse desmotivados ya sea por enfrentar una clase muy difícil o demasiado fácil, por lo que pueden abandonar el curso o al final de cuentas tener que repetirlo. En los casos en que se dan cuenta del problema y fueron ubicados

en un nivel inferior, pueden llegar a distraer a los demás estudiantes al no poner la atención debida en la clase. En cualquiera de los casos los alumnos pueden sentirse frustrados e inclusive molestos ante la pérdida de tiempo e incluso, de dinero. Las inconveniencias administrativas pueden ser muchas al tratar de reubicar a los alumnos en la clase adecuada, que en ocasiones puede no estar disponible, además de la pérdida de tiempo y la inversión en costo por el trabajo extra del personal encargado de estos asuntos. A nivel programa puede ser problemático para el profesor y los alumnos que, después de varios días o incluso semanas de clase en la que ya se ha alcanzado cierto avance, ritmo e integración entre los participantes, llegue un estudiante nuevo quien tiene que actualizarse con lo ya estudiado, así como pasar por un nuevo proceso de adaptación.

Por lo tanto, un proceso de colocación adecuado, preciso y pertinente que cuente con los instrumentos adecuados y el personal académico y administrativo necesario, ayudará a facilitar que el proceso de enseñanza-aprendizaje de una lengua extranjera sea exitoso.

▪ **Preguntas de investigación**

Desde hace varios años el Centro de Idiomas de la Universidad Iberoamericana cuenta con un procedimiento para establecer el nivel de conocimiento del idioma inglés y así poder ubicar a los interesados en el curso que les permita mejorar sus habilidades lingüísticas y adquirir de manera gradual una mayor competencia en la lengua meta.

De acuerdo con la problemática general mencionada anteriormente con respecto al proceso de clasificación, surgen las siguientes preguntas de investigación:

Pregunta genérica:

- ¿Cuál es la efectividad del proceso de colocación que se lleva a cabo en el Centro de Idiomas de la UIA en cuanto a su confiabilidad y evidencia de validez?

Para poder dar respuesta a la cuestión anterior necesitaremos responder las siguientes preguntas específicas:

- En un semestre de instrucción ¿existen alumnos que tengan que ser reubicados en niveles distintos al inicial?
- ¿En cuanto tiempo se determina que el estudiante ha sido mal clasificado y quién se percata de ello (el estudiante, el profesor, la coordinación)?
- ¿Cuál es la percepción que tienen los usuarios (alumnos, examinadores y profesores) del proceso de selección en su conjunto con respecto a su utilidad, confiabilidad y validez?
- ¿Existe algún elemento del proceso (examen, entrevista, examinadores, etc.) que pudiera contribuir a la existencia de errores en la clasificación?

▪ **Objetivo**

Analizar la confiabilidad y evidencia de validez del proceso para determinar el nivel de inglés de los aspirantes a ingresar a los cursos que ofrece el Centro de Idiomas de la Universidad Iberoamericana.

▪ **Justificación**

La relevancia social de este proyecto está relacionada con el proceso de la globalización y los cambios que ha traído consigo en los aspectos económicos, políticos, sociales, culturales y educativos, los cuales han dado pie tanto a oportunidades como a nuevos retos para los distintos países que, voluntariamente o de manera forzada, se han integrado a dicho proceso.

En el caso de la Educación Superior en México este fenómeno también ha producido nuevos enfoques para el desarrollo y necesidades por atender en este nivel de educación. Entre ellos mencionaremos el que se refiere a ofrecer estudios superiores de calidad que permitan a los

egresados “solventar las necesidades humanas de corte ético [y] enfrentar un sistema mundial crecientemente complejo.” (Gacel, 1998).

Hoy en día se espera que las instituciones de educación superior como son las universidades, ofrezcan programas de calidad pertinentes, pues “deben desempeñar un papel crucial en el sostenimiento de la dignidad humana y en la resolución de los problemas globales.” (Ibid) Asimismo, se considera que ésta es una de las responsabilidades de mayor relevancia para dichas instituciones pues de no ser así “estamos haciendo un grave perjuicio a nuestros estudiantes si no los preparamos adecuadamente para vivir y trabajar en un mundo complejo, interdependiente y multicultural.” (Johnston en Gacel, 1998)

Por lo anteriormente expuesto, el aspecto fundamental en torno a la importancia social de este proyecto está basado en la necesidad de contribuir con una educación superior de calidad, al menos para quienes, sorteando las vicisitudes generadas por el contexto global en el que vivimos, logran llegar a este nivel educativo.

Son muchos los elementos que sin duda conforman lo que podríamos llamar una educación superior de calidad en este ya no tan nuevo contexto de una *aldea global*. Sin embargo, son dos los puntos que mencionaré en esta justificación. Por un lado, la inminente necesidad de la internacionalización de las instituciones de educación superior y, dentro de este marco, la importancia del aprendizaje de idiomas, y en particular del inglés.

La *internacionalización educativa* es descrita por Knight como “...una de las maneras por la cual un país responde al impacto de la globalización, respetando la individualidad de cada nación...” (en Gacel, 2005). Dentro de las instituciones educativas se considera como “...un proceso de transformación institucional que tiene [como] estrategia la integración de la dimensión internacional e intercultural en la misión, cultura, planes de desarrollo y políticas generales de las Instituciones de Educación Superior (IES).” (Ibid)

En suma, las IES han de procurar que sus programas de formación permitan a sus egresados vivir, comprender y respetar la interculturalidad tan característica del mundo actual. Esto habrán de hacerlo sin dejar de lado la orientación humanista, la cual es necesaria para lograr el desarrollo económico con equidad, tan necesario de fomentar en un país tan desigual como el nuestro.

¿Cómo lograr una dimensión internacional en las IES? De acuerdo con Gacel (2005) los mencionados programas de internacionalización que las IES deben incluir en su misión universitaria deben llevarse a cabo a través de un desarrollo institucional que abarque todos los ámbitos de la institución, como lo son los aspectos culturales, académicos y administrativos que conforman las bases de una comunidad universitaria.

Es en los dos primeros ámbitos, cultural y académico, donde podemos ubicar en gran medida la importancia del aprendizaje de idiomas. Al referirse a este punto en particular, Gacel (2005) indica que “la internacionalización implica el acercamiento a otras culturas y, en consecuencia, a otros idiomas. La falta de dominio de un segundo o tercer idioma limita el aprovechamiento de oportunidades de cooperación y programas de becas e intercambio con instituciones y organismos extranjeros tanto para el personal académico y administrativo como para los estudiantes.” Y como ella misma establece, es importante que las instituciones de educación superior en México determinen como prioritaria y obligatoria la enseñanza de idiomas extranjeros a través de programas académicos de calidad.

La necesidad de internacionalizar a las universidades nos conduce a un área hasta cierto punto poco considerada en la realidad de México como factor de importancia en el sistema educativo: la enseñanza de lenguas extranjeras para lograr una comunicación efectiva entre países, mercados, empresas e individuos.

¿Cuáles son las ventajas? El aprendizaje de lenguas extranjeras no sólo incide en las relaciones interpersonales porque genera el conocimiento y respeto por otras culturas, elemento imprescindible para que las experiencias que se establecen en cualquier ámbito, ya sea político, de negocios o social, sean experiencias con mayores posibilidades de éxito.

Adicionalmente, el conocimiento de un idioma extranjero otorga al individuo valiosas ventajas como lo es conocer mejor no sólo su idioma materno, sino la función de las lenguas en general. Cuando desarrollamos un dominio sobre otra lengua como sistema lingüístico y adquirimos conocimientos sobre su fonología, morfología y sintaxis, así como sobre la manera en que se utiliza en determinado contexto sociocultural, pasamos por un proceso mediante el cual llegamos a comprender la naturaleza del lenguaje mismo. Esto es, desarrollamos un sentido sobre la estructura, que es difícil de adquirir cuando aprendemos nuestra propia lengua. Es así como las competencias adquiridas en una lengua se transfieren a otra para reforzar su comprensión.

Además de ello, la adquisición o el aprendizaje de otros idiomas estimula el desarrollo de habilidades cognitivas, es decir, favorece la capacidad para aprender en general pues “algunos estudios en educación bilingüe parecen llevar a la conclusión de que el estudio de lenguas extranjeras puede ayudar y hasta acelerar el desarrollo cognitivo del cerebro. Los sujetos bilingües en diversas pruebas sacaron puntuaciones significativamente más altas que los sujetos monolingües en pruebas verbales y no verbales de inteligencia.” (Barrón et al, 2002). Estos autores indican que también favorece el desarrollo de la competencia comunicativa es decir, el interactuar de manera efectiva para intercambiar información en forma verbal, escrita y gráfica, por medios interpersonales, masivos y electrónicos, en diferentes contextos y utilizando diferentes códigos tales como lengua materna, segunda lengua, mensaje no verbal, etc.

Ahora bien, la internacionalización de México en el mundo actual nos lleva a la necesidad, como ya se ha mencionado, de reforzar los aspectos académico, profesional y personal de los

individuos. En el caso de los estudiantes que ingresan al nivel superior en un principio simplemente se espera que puedan realizar ciertas lecturas académicas en otros idiomas. Para cuando terminan sus estudios es necesario que tengan un nivel de comunicación eficiente en caso de verse enfrentados a dicha necesidad. En conclusión, contar con las competencias de comprensión y producción en otra lengua, puede repercutir en un desempeño laboral más ventajoso ya sea en México o, incluso, en el extranjero.

En este trabajo nos enfocamos al idioma inglés que es considerado desde hace décadas como *el idioma universal*. De acuerdo con Altbach (2004), el inglés “es el latín del siglo XXI”. El dominio del inglés no es sorprendente tomando en cuenta que Estados Unidos ha sido considerado una potencia mundial. Y qué decir de otro país destacado dentro de las naciones desarrolladas, como lo es Inglaterra. El idioma inglés es tan importante en la globalización, que de acuerdo con Cristal (en Albatch, 2004) merece un análisis detallado en virtud de que es un aspecto con el que la educación mundial debe lidiar.

La realidad es que de una u otra forma, el inglés es el idioma más estudiado no sólo como segunda lengua, sino también como lengua extranjera. La razón fundamental es que se utiliza continuamente en revistas científicas, educativas, económicas, etc., así como en ponencias a nivel internacional. Todo el mundo asume que “si está en inglés” un documento de cualquier índole llegará a más personas.

Por lo tanto, al destacar la importancia de la internacionalización de las universidades, la relevancia del manejo del idioma inglés, y al retomar el planteamiento del problema del presente proyecto, donde he mencionado el caso del Centro de Idiomas de la UIA, debo enfatizar que la importancia social de este proyecto radica en contribuir a una educación internacional de calidad. Para lograrlo es necesario poder ofrecer a la comunidad universitaria no sólo un programa pertinente para el aprendizaje de este idioma, sino un comienzo sólido al iniciar dicho programa.

Es decir, un proceso de evaluación para determinar el nivel de inglés de los aspirantes a los cursos que ofrece la UIA, que sea de calidad y cuente con un buen grado de validez y confiabilidad.

El presente estudio tiene como finalidad determinar la efectividad actual de estos elementos, y en su caso, hacer las recomendaciones pertinentes para mejorarlo.

II. MARCO REFERENCIAL

▪ **Aspectos lingüísticos²**

- El lenguaje

La primera pregunta que debemos responder para iniciar esta sección es ¿qué es el lenguaje?

De acuerdo con el Diccionario de Lingüística Moderna (1997), el significado más general del lenguaje es “la capacidad que tienen los seres humanos para comunicarse entre sí por medio de signos lingüísticos, primariamente en el medio oral-auditivo y, secundariamente, en el medio visual o en el háptico.”³

El medio oral-auditivo está formado por el aire, que es el medio físico que transmite los sonidos articulados, y forma estructuras (modelos) de rápida producción, efímeras en el tiempo. En el medio visual estas estructuras son más lentas, elaboradas y duraderas.

Pero el lenguaje es mucho más que un sistema de sonidos y modelos. El lenguaje es un sistema de comunicación con distintos niveles de complejidad que incluye la difícil selección y ordenamiento de estructuras con significado. Este significado es arbitrario dependiendo del idioma específico, sin embargo es compartido por una misma comunidad de habla, es decir, por personas que poseen una misma cultura, y que por lo tanto se comunican a través de los mismos significados culturales.

Además del significado cultural existe un significado individual, el cual hace uso de explicaciones, metáforas, analogías, y otras formas indirectas que se transmiten a través del

² La Lingüística es la ciencia del lenguaje. En este sentido el término lenguaje comprende la facultad comunicativa propia de los humanos y las lenguas naturales en las que éste se manifiesta, junto con sus elementos constitutivos, las relaciones entre éstos, y las leyes o reglas de su funcionamiento, de describir y explicar sus formas y funciones como expresión del pensamiento humano y de la comunicación social. (Aarens, 1976; Robins, 1980; Sampson, 1980 en Alcaraz y Martínez, 1997).

³ El medio háptico (o táctil) es el empleado por el sistema de escritura Braille.

lenguaje. Cuando usamos el lenguaje de manera individual como nativos hablantes de cualquier idioma, no nos percatamos de que estamos utilizando una herramienta sumamente compleja. Por ejemplo, seguimos la línea de una conversación decidiendo qué decir, qué preguntar; le decimos a nuestro interlocutor qué hacer, elegimos las palabras específicas para cierta situación, y para la persona específica que nos escucha; elegimos oraciones, sufijos y prefijos, conectamos cada oración y cada parte de la oración con lo que se dijo antes o se dirá después; acentuamos palabras, modulamos el tono de voz para indicar significado, dependiendo de nuestro estado de ánimo. Hacemos todo esto tomando en cuenta el efecto que tendrá en nuestro oyente, al tratar de agradarlo, alabarlo, reprenderlo; hacerle alguna sugerencia o petición, o darle alguna orden, etc., y hacemos todo esto en un promedio de cientos de movimientos de nuestras articulaciones⁴. (Lado, 1964)

El manejo del lenguaje implica dominar la información que se encuentra en nuestra memoria y que está lista para usarse a través de estrategias metacognitivas con las cuales podemos interpretar el discurso. Para Bachman y Palmer (1997), existen dos amplias categorías del manejo del lenguaje, a las que llaman *conocimiento organizacional*, y *conocimiento pragmático*. La primera se refiere al manejo de las estructuras formales del lenguaje para producir y comprender expresiones y oraciones gramaticalmente correctas para poder crear textos tanto orales como escritos. La segunda nos permite crear o interpretar el discurso al relacionar las expresiones, oraciones y tipos de texto con sus significados, con la intención de quien las usa, y con las características relevantes del contexto en que se aplican

Conocer cuáles son las habilidades del lenguaje es fundamental para el desarrollo y uso de exámenes de lengua. Para Bachman (1990) poseer dichas habilidades no significa exclusivamente saber hablar, escribir (producir), leer y escuchar (comprender), ni dominar sus respectivos

⁴ En fonética, el proceso que sufren los órganos de la cavidad bucal (faringe, fosas nasales, tracto vocálico, labios, paladar, etc.) para emitir sonidos.

componentes como la gramática, el vocabulario y la pronunciación. Por el contrario, poseer las habilidades del lenguaje significa poder ir más allá de la oración para poder hacer uso apropiado de la lengua meta. Esto es, hacemos uso del lenguaje de manera contextual, puesto que el lenguaje no se utiliza de forma aislada (no sólo ‘escuchamos’, sino que escuchamos algo específico, con un propósito y contexto particular).

Como sugiere Lado (1964), cada idioma es un sistema comunicativo en el que los sonidos, palabras y estructuras sintácticas adquieren significado con base en las condiciones particulares de dicho sistema y la experiencia y contexto cultural de los hablantes de esa lengua.

- Aprendizaje vs. Adquisición, una base para la elaboración de exámenes.

Los exámenes de lengua son valiosas fuentes de información que nos ayudan para diversos propósitos en el desarrollo de la enseñanza-aprendizaje. A través de ellos podemos diagnosticar fortalezas y debilidades de un estudiante, determinar su grado de progreso, en dónde iniciar un programa de estudios, decidir si un alumno está listo para un cambio de grado, etc. Es decir, nos ayudan a tomar decisiones importantes que pueden beneficiar (o afectar) a un individuo, por lo que para poder elaborar y/o evaluar exámenes de lengua que sean adecuados, precisos, válidos y confiables, debemos tener muy claro el nivel de competencia y dominio lingüístico que esperamos posea un individuo en determinado momento y para determinado fin.

Para ello, necesitamos considerar el contexto en el que se enseña una lengua distinta a la materna, y distinguir la diferencia entre adquisición de una segunda lengua (L2) y aprendizaje de una lengua extranjera (LE).

Adquisición se refiere al “proceso en el cual el individuo adquiere la lengua de manera *natural*, intuitiva e inconsciente, sin poner atención a la forma lingüística”, y aprendizaje se

entiende como “un proceso de desarrollo consciente de la lengua” (Krashen en Da Silva y Signorete, 2005)

Al respecto coincido con Da Silva y Signorete, cuando afirman que: “La distinción entre L2 y LE no siempre está señalada en los textos especializados, algunos teóricos piensan que estos términos no expresan fenómenos significativamente distintos. Sin embargo [...] sus procesos de desarrollos no son iguales[...] las variables afectivas influyen de manera distinta en cada uno de esos dos desarrollos y [...] el tipo de *input* o información lingüística que se recibe para lograr el dominio del comportamiento verbal en L2 y LE es generalmente muy diferente. [...] L2 se desarrolla en situaciones similares a las de la adquisición de L1, mientras que LE es aprendida en las clases de lenguas”.

Es decir, un nativo-hablante de español podría llegar a adquirir una L2 siempre y cuando esté inmerso en el contexto de la lengua meta durante un considerable periodo de tiempo (y, de acuerdo con la Teoría del Periodo Crítico, de preferencia a una edad temprana⁵). Al estar expuesto las veinticuatro horas del día a la radio, la televisión y a la interacción con nativo-hablantes, sería posible que la adquiriera de manera inconsciente. En algunos casos excepcionales cierto grado de adquisición pudiera presentarse aún sin estar inmerso en el país donde se habla el idioma de interés, como es el caso de algunas escuelas privadas consideradas 100% bilingües y que enseñan el idioma desde la educación preescolar. En estos contextos la elaboración de los exámenes de lengua puede basarse en el uso cotidiano del lenguaje, y por lo general pueden ser respondidos de manera intuitiva, sin un análisis necesariamente profundo de las estructuras.

Aprendemos un idioma como LE cuando, como en México, recibimos la instrucción formal de una lengua extranjera en las escuelas como parte del sistema educativo. En el caso de nuestro

⁵ Esta teoría se refiere a que hay una edad después de la cual nuestra habilidad de aprender una lengua se limita significativamente. Existe controversia con respecto a cuál es exactamente esta edad crítica, pero en general se considera que ésta se da alrededor de los 12 años es decir al inicio de la pubertad. Tal parece entonces que la infancia temprana y media es el periodo de la vida de un individuo en la que éste se encuentra mejor “equipado” para aprender una lengua. (Anula, 1998).

país, esto ocurre a partir del nivel secundaria en las escuelas públicas⁶. En ocasiones, debido al limitado conocimiento aprendido en estas instituciones, se recurre a alguna escuela de enseñanza especializada en el idioma de interés. Bajo este contexto, los exámenes de lengua tienden a basarse en el tipo de problemas que un nativo-hablante de otro idioma enfrenta en el uso de la lengua meta, por lo que el análisis por parte del estudiante será más profundo y basado en las estructuras y reglas aprendidas.

De acuerdo con el Marco Común Europeo de Referencia para las Lenguas (MCERL, 2002), en la actualidad no existe un consenso con respecto a la manera en que los alumnos no nativos aprenden una lengua específica como para determinar una teoría del aprendizaje concreta. Algunos teóricos consideran que las capacidades humanas de procesamiento de información son lo suficientemente fuertes como para que una persona que se vea expuesta a una lengua comprensible pueda aprenderla y logre utilizarla tanto en el ámbito de comprensión como de producción.

El MCERL afirma que otra postura al respecto indica que además de que el estudiante sea expuesto de manera comprensible al material de entrada (*input*), una condición necesaria para el desarrollo de la lengua es la participación activa y la interacción comunicativa. Además menciona otro punto de vista en el que se cree que el aprender las reglas gramaticales necesarias y el vocabulario suficiente permitirá al estudiante comprender y utilizar la lengua con base en su sentido común y en experiencias previas.

El complejo proceso de comunicación que se logra a través del lenguaje es posible gracias a un sistema de hábitos que operan en gran medida de manera inconsciente. Cuando se adquiere la L1 ó la L2, estos hábitos están profundamente instalados en el sistema nervioso y en el proceso intelectual y emocional de un individuo. Cambiar cualquier parte de este sistema de hábitos sería

⁶ Posible razón por la que, de acuerdo con la teoría del periodo crítico anteriormente comentada, los egresados de estas instituciones presenten serias dificultades con el manejo del idioma.

una ardua empresa, como lo sería lograr establecer el mismo sistema para el aprendizaje de una lengua extranjera.

Hemos dicho que cuando el nativo hablante de una lengua expresa algo a su interlocutor, ya sea nativo hablante o alguien que haya *adquirido* el idioma, emplea una secuencia cultural de significados lingüísticos, de oraciones, frases, palabras, morfemas y fonemas. Lo hace sin pensar, y la persona que lo escucha reacciona a ello de la misma manera inconsciente y a la misma velocidad.

En el caso de un individuo con otro sistema lingüístico o idioma, que haya *aprendido* la lengua, ocurre lo contrario. Al escuchar a su interlocutor no puede evitar utilizar primero su propio sistema de hábitos, el cual no puede ser eliminado a voluntad. Por lo tanto en este caso, escucha primero unidades de sonido, palabras, frases y oraciones y los transfiere a su propio idioma. Esto es, distorsiona lo que escucha para adecuarlo a su sistema lingüístico. Asimismo cuando intenta expresarse, piensa en el significado general de lo que quiere decir en su propio código y lo transforma en las unidades de lenguaje mencionadas, sin percatarse con facilidad de posibles errores en la lengua meta, y aún cuando se le indique que ha cometido un error no siempre logrará comprender la causa o le llevará cierto tiempo lograr corregirlo.

Es por esto que, como ya se mencionó, reconocer los problemas, es decir, las estructuras que difieren entre dos idiomas, es una manera de aprender y evaluar una lengua extranjera. Reconocer o detectar las diferencias que se presentan entre una lengua y otra es, en principio, conocerla, además de que será la base de las evaluaciones o exámenes, los cuales serán válidos y confiables dependiendo de su propio objetivo de medición. Las unidades y patrones que no tienen una contraparte en el lenguaje materno o que la tienen pero con una distribución y significado estructuralmente diferentes, nos permitirán elegir los modelos a evaluar en los exámenes de lengua extranjera.

- La competencia y el dominio lingüístico

De acuerdo con el MCERL (2002), la finalidad de la enseñanza de idiomas es lograr que los estudiantes manejen un nivel de competencia y dominio en una lengua determinada. Desde un enfoque intercultural, el objetivo de la educación en una lengua es promover en el individuo el desarrollo personal así como un sentido de identidad como respuesta a la enriquecedora experiencia que como agente social implica enfrentarse a lo diferente en los ámbitos de la lengua y la cultura.

El uso de la lengua –que incluye el aprendizaje- comprende las acciones que realizan las personas que, como individuos y como agentes sociales, desarrollan una serie de competencias, tanto generales, como competencias comunicativas lingüísticas, en particular (MCERL, 2002).

Las *competencias* son la suma de conocimientos, destrezas y características individuales que permiten a una persona realizar acciones. Por otro lado, las *competencias generales* son aquellas que no se relacionan directamente con la lengua, pero a las que se puede recurrir para acciones de todo tipo, incluyendo actividades lingüísticas. Por último la *competencia comunicativa* es la que permite a una persona actuar utilizando específicamente medios lingüísticos. Estas dimensiones se encuentran relacionadas entre sí y son parte del uso y aprendizaje de una lengua (MCERL, 2002).

Las clases de lengua deben fomentar el desarrollo de la competencia comunicativa, la cual está compuesta por tres factores: el *lingüístico*, que incluye conocimientos y destrezas léxicas, fonológicas y sintácticas; el *sociolingüístico*, que se refiere a las condiciones socioculturales del uso de la lengua e incluye el respeto a las convenciones sociales (como normas de cortesía, normas entre generaciones, sexos, clases, grupos sociales, etc); y por último el *pragmático*, que tiene que ver con el uso funcional de los recursos lingüísticos y con el dominio del discurso, la

cohesión, la coherencia, los tipos de texto, la ironía, etc., el cual ejerce un fuerte impacto en las interacciones y entornos culturales (MCERL, 2002).

El MCERL establece una serie ascendente de niveles comunes de referencia para describir el dominio que un alumno tiene de la lengua. Estas descripciones contribuyen a concretar lo que se espera sea adecuado en distintos niveles de aprovechamiento y contribuyen al desarrollo de especificaciones con respecto a los objetivos generales de aprendizaje de manera flexible.⁷

A continuación se presenta la escala global y las competencias esperadas en los niveles de dominio que el MCERL establece:

Tabla 1. Escala de dominio lingüístico

Usuario competente	C2	Es capaz de comprender con facilidad prácticamente todo lo que oye o lee. Sabe reconstruir la información y los argumentos procedentes de diversas fuentes, ya sean en lengua hablada o escrita, y presentarlos de manera coherente y resumida. Puede expresarse espontáneamente, con gran fluidez y con un grado de precisión que le permite diferenciar pequeños matices de significado incluso en situaciones de mayor complejidad.
	C1	Es capaz de comprender una amplia variedad de textos extensos y con cierto nivel de exigencia, así como reconocer en ellos sentidos implícitos. Sabe expresarse de forma fluida y espontánea sin muestras muy evidentes de esfuerzo para encontrar la expresión adecuada. Puede hacer un uso flexible y efectivo del idioma para fines sociales, académicos y profesionales. Puede producir textos claros, bien estructurados y detallados sobre temas de cierta complejidad, mostrando un uso correcto de los mecanismos de organización, articulación y cohesión del texto.
Usuario independiente	B2	Es capaz de entender las ideas principales de textos complejos que traten de temas tanto concretos como abstractos incluso si son de carácter técnico, siempre que estén dentro de su campo de especialización. Puede relacionarse con hablantes nativos con un grado suficiente de fluidez y naturalidad, de modo que la comunicación se realice sin esfuerzo por parte de los interlocutores. Puede producir textos claros y detallados sobre temas diversos, así como defender un punto de vista sobre temas generales, indicando los pros y los contras de las distintas opciones.
	B1	Es capaz de comprender los puntos principales de textos y en lengua estándar si tratan sobre cuestiones que le son conocidas, ya sea en situaciones de trabajo, de estudio o de ocio. Sabe desenvolverse en la mayor parte de las situaciones que pueden surgir durante un viaje por zonas donde se utiliza la lengua. Es capaz de producir textos sencillos y coherentes sobre temas que le son familiares o en los que tiene un interés personal. Puede describir experiencias, acontecimientos, deseos y aspiraciones, así como justificar brevemente sus opiniones o explicar sus planes.
Usuario básico	A2	Es capaz de comprender frases y expresiones de uso frecuente relacionadas con áreas de experiencia que le son especialmente relevantes (información básica sobre sí mismo y su familia, compras, lugares de interés, ocupaciones, etc.). Sabe comunicarse a la hora de llevar a cabo tareas simples y cotidianas que no requieran más que intercambios sencillos y directos de información sobre cuestiones que le son conocidas o habituales. Sabe describir en términos sencillos aspectos de su pasado y su entorno, así como cuestiones relacionadas con sus necesidades inmediatas.

⁷ Cabe aclarar que, como su nombre lo indica, el Marco sirve de referencia, pues no pretende decir a los profesionales lo que tienen que hacer o de qué forma hacerlo (MCERL, 2002). Sin embargo, este marco ha sido aceptado y adoptado por diversos países e instituciones educativas.

A1	<p>Es capaz de comprender y utilizar expresiones cotidianas de uso muy frecuente, así como frases sencillas destinadas a satisfacer necesidades de tipo inmediato.</p> <p>Puede presentarse a sí mismo y a otros, pedir y dar información personal básica sobre su domicilio, sus pertenencias y las personas que conoce.</p> <p>Puede relacionarse de forma elemental siempre que su interlocutor hable despacio y con claridad y esté dispuesto a cooperar.</p>
----	---

Tomado del MCERL (2002)

▪ Evaluación educativa

Además de que la evaluación educativa es considerada cada vez más una actividad que favorece la calidad y mejora de la enseñanza, la sociedad en general así como los educadores, padres de familia y alumnos en particular, están cada vez más conscientes de lo que significa evaluar o ser evaluado y las consecuencias positivas o negativas que la toma de decisiones basadas en ello pueden traer consigo. Es decir, existe cada vez más una *cultura de la evaluación* (Castillo, 2003).

A pesar de su reciente auge, es necesario destacar que el concepto y objetivos de la evaluación educativa han ido cambiando a lo largo de su historia y así, su importancia. Castillo (2003) establece 6 periodos de la evaluación como sigue:

1. De los inicios del S XX a los años 20 se consideraba la evaluación como *medida* y se basaba en la Psicología Conductista (Skinner). Poco tenía que ver con programas de estudio.
2. En los años 30 y 40 la evaluación estaba relacionada con el grado de consecución de los objetivos propuestos en un programa. Tyler concibió a la educación como un proceso sistemático destinado a producir cambios en la conducta de los alumnos, y a la evaluación como el mecanismo que permitía comprobar el grado de consecución de los objetivos propuestos.
3. Hasta finales de los 60 se le percibió como la totalidad del ámbito educativo. Los autores que más contribuyeron a este concepto fueron Cronbach, quien la define como

la recopilación y uso de información para la toma de decisiones, y Scriven como el proceso por el cual se estima el mérito o valor de algo que se evalúa.

4. Durante la década de los 70 la evaluación se orientaba a dos niveles: los alumnos y la toma de decisiones sobre el programa correspondiente, y se entendía como la valoración del cambio ocurrido en el alumno como consecuencia de una acción educativa sistemática. En este tiempo destacan las taxonomías de Bloom, Mager y Gagné.
5. En los 80 surgen nuevos modelos relacionados con la evaluación cuantitativa y cualitativa.
6. Es a partir de los 90 cuando la evaluación se vuelve globalizada e integradora.

La contribución de quienes se dedicaron y siguen dedicando a la evaluación ha sido de gran valor para sentar las bases de lo que *evaluar* significa en nuestros días. Algunas de las definiciones que surgieron a partir de los 90 incluyen las siguientes (tomadas de Castillo 2003, excepto la que se indica)⁸ :

- Benedito, 1990: “Evaluar es una actividad sistemática, continua e integrada en el proceso educativo, cuya finalidad es conocer y mejorar al alumno en particular y al proceso educativo, con todos sus componentes, en general.
- Casanova 1995: “Proceso sistemático y riguroso de recogida de datos, incorporado al proceso educativo desde su comienzo de manera que sea posible disponer de información continua y significativa para conocer la situación, formar juicios de valor con respecto a ella y tomar las decisiones adecuadas para proseguir la actividad educativa, mejorándola progresivamente”

⁸ Castillo (2003, págs. 6-7) ofrece una detallada recopilación de las diversas definiciones de autores reconocidos en el campo y la historia de la evaluación.

- Para la Junta de Andalucía (1996), evaluar significa “enjuiciar, tomar decisiones sobre nuevas acciones a emprender y, en definitiva, transformar para mejorar.”
- Rodríguez Diéguez 1998: “La evaluación consiste en el proceso y resultado de la recogida de información sobre un alumno o un grupo de clase con la finalidad de tomar decisiones que afecten a las situaciones de enseñanza”
- Castillo Arredondo 2002: “La evaluación debe permitir, por un lado, adaptar la actuación educativo-docente a las características individuales de los alumnos a lo largo de su proceso de aprendizaje; y por otro, comprobar y determinar si éstos han conseguido las finalidades y metas educativas que son el objeto y la razón de ser de la actuación educativa.”

El nuevo concepto concebido a partir de los noventa hace que la evaluación se convierta en un proceso sistemático y riguroso de recogida de datos con el fin de contar a lo largo del proceso educativo de información pertinente para formar juicios de valor y tomar decisiones acertadas en torno a la actividad educativa para así mejorarla de manera continua.

El cambio en el concepto es fácil de reconocer si recordamos que anteriormente la evaluación estaba relacionada únicamente con la aplicación de un *examen*, a través del cual se determinaba si el alumno manejaba o no los objetivos marcados en un programa de estudio, lo que determinaba si el estudiante *aprobaba* o *reprobaba*. Es decir, se solía utilizar la medición para calificar, seleccionar y controlar, muchas veces de manera descontextualizada y sin sentido.

Hoy en día *evaluar* incluye todo el proceso educativo: desde los programas de estudio, el procedimiento de ingreso, pasando por las actividades en clase, el profesor, la administración, etc., y el examen se ha vuelto simplemente un instrumento valioso para la recogida de información. Un gran número de instituciones y de docentes intentan hacer suyo este nuevo

concepto, y aunque aún se presenta cierta disparidad de criterios, su aplicación se ha ido fomentando de manera gradual (Elola y Toranzos, 2000).

De acuerdo con Castillo (2002), la evaluación abarca los siguientes ámbitos del sistema educativo:

- Los procesos de enseñanza
- Los procesos de aprendizaje
- La práctica docente
- Los proyectos curriculares
- Los diversos elementos del sistema educativo

- La evaluación del proceso enseñanza-aprendizaje

La finalidad última de la evaluación es el seguimiento del proceso de enseñanza-aprendizaje con el fin de mejorarlo, reconducirlo y adecuarlo de manera continua.

La estructura básica del concepto de evaluación como proceso dinámico, abierto y contextualizado, incluye tres pasos fundamentales que son:

1. Obtener información
2. Formular juicios
3. Tomar decisiones

Además, existen diversos tipos de evaluación que se distinguen en función de los diversos ámbitos y circunstancias de su aplicación. Se distinguen tres momentos:

- Antes del proceso E-A, es decir la *evaluación inicial* que puede ser diagnóstica, pronóstica o previsor.

- Durante el proceso E-A, conocida como *evaluación formativa*, cuyo objetivo es ser orientadora, reguladora y motivadora.
- Después del proceso de E-A, llamada *evaluación sumativa*, que debe ser integradora, promocional y acreditativa.

- **Medición del aprendizaje**

Es necesario distinguir la diferencia entre evaluación y medición del aprendizaje. La primera supone hacer juicios de valor con respecto al desempeño del estudiante para tomar decisiones. Lo hacemos con base en ciertos valores y comparamos la información con criterios específicos previamente establecidos.

La medición es la descripción numérica de un suceso o una característica. La medición del aprendizaje tiene por objetivo informar *cuánto*, o *qué tan frecuentemente* mediante puntuaciones o rangos. La medición del aprendizaje tiene una función importante para la toma de decisiones, pero su resultado no puede considerarse como absoluto. Es decir, la medición es diferente a la evaluación porque no proporciona juicios de valor. Sirve de base para la evaluación (aunque no todas las evaluaciones requieran de una medición), pero es tan sólo una etapa de ésta.

La medición del aprendizaje puede ser compleja porque éste no tiene nunca un valor cero (esto es, algo se sabe, por poco que esto sea), el cual representa el punto de partida de cualquier medición. Además, es difícil tener un patrón válido de comparación. Frecuentemente este patrón es el número de preguntas de la prueba y con esa base se efectúa la medición.

Ahora bien, cuando pretendemos medir algo y utilizamos algún instrumento de medida para hacerlo, debemos asegurar que éste cumpla con dos criterios fundamentales de la evaluación: la validez y la confiabilidad.⁹

- Confiabilidad

Uno de los elementos imprescindibles de la evaluación y de la medición es la confiabilidad y se refiere a la medida en que los resultados de un procedimiento o prueba evaluada son fiables. La pregunta que surge cuando queremos determinar el grado de confiabilidad es: si los candidatos o estudiantes volvieran a pasar por el mismo proceso, o a responder la misma prueba un tiempo después de haberla hecho hoy en condiciones similares a las de la primera aplicación, ¿obtendrían el mismo resultado (asumiendo que la cualidad medida no hubiera cambiado)? En otras palabras, nos referimos a la capacidad de la prueba para demostrar consistencia y estabilidad en las puntuaciones.

De las diversas definiciones de confiabilidad mencionadas en Castillo (2002), destacamos las siguientes:

- “Es una característica fundamental de los instrumentos de medida. Un instrumento es fiable si, utilizado dos o más veces en circunstancias similares, se obtienen los mismos resultados. La fiabilidad es un requisito imprescindible en toda medición, pero no suficiente” (López-Barajas, 1996).

La siguiente la elegimos por referirse en general a un proceso y no exclusivamente a las pruebas o exámenes.

⁹ Un criterio adicional que algunos autores señalan como importantes, es el aspecto práctico (o factibilidad) del instrumento, es decir, su conveniencia, costo, facilidad de corrección y aplicación (Karmel, 1984).

- “Requerimiento metodológico que exige que el fenómeno objeto de estudio pueda ser observado repetidamente. Se refiere al grado de estabilidad de un procedimiento o medida” (Corral, 1997).

Por su parte, Alderson, et al (1998) comentan que es casi imposible obtener una prueba absolutamente confiable, pero recomiendan garantizar que el proceso o instrumento se administre y, en su caso, se corrija, de manera consistente; que las instrucciones sean claras, y en el caso de los exámenes, que no haya ítems ambiguos o defectuosos, lo que permitirá mayor confiabilidad.

Indican también que la confiabilidad de una prueba depende entre otras cosas del tipo y longitud de ésta y, al referirse a exámenes de lengua, de la gama de competencia lingüística de los alumnos.

En el caso de las pruebas objetivas (esto es, aquellas que utilizan ítems del tipo de respuesta de opción múltiple, verdadero-falso, y reconocimiento del error, entre otras, en las que el candidato elige la respuesta que se considera como “correcta” o “incorrecta”), la confiabilidad depende de la homogeneidad de los ítems, así como de otros factores que pueden provocar inconsistencias, como la diferencia en la condición de los individuos que serán examinados en cuanto a su estado de ánimo, o físico; diferencias en el contenido de los instrumentos en el caso de que se apliquen dos versiones; variaciones en la administración de la prueba, como ruido o en la manera en que el examinador lo aplica; así como errores en la calificación y en el registro de las puntuaciones (Karmel, 1984).

En cuanto a las pruebas subjetivas (en las que los examinadores no pueden aplicar juicios del tipo “correcto” “incorrecto”, sino que deben evaluar la calidad de la respuesta de un candidato a una tarea con base en una “escala de valores”, como sucede en los exámenes orales utilizados en las evaluaciones lingüísticas, o en los ensayos), Alderson, et al (1998) indican que la confiabilidad es más difícil de obtener porque existen diversos factores que pueden afectarla.

Algunos de ellos pueden ser la variación en la aplicación del instrumento, la ambigüedad de los criterios de evaluación; la aplicación de distintos modelos por distintos evaluadores, y la inconsistencia por parte de algunos de ellos en su toma de decisiones. Sin embargo afirman que existen procedimientos para asegurar que los examinadores logren utilizar los mismos criterios.

- **Validez**

El concepto de validez ha sufrido distintas transformaciones a lo largo de los años. Las más antiguas definiciones la consideraban un elemento estático de un examen o instrumento de medida. Por ejemplo, la establecida por Guilford (1946, en Goodwin, 2002) que en términos generales afirmaba que una prueba era válida para cualquier cosa con la que existiera una correlación, lo que significa simplemente que un examen podía ser válido, o no serlo.

En 1966, se creó un comité formado por la Asociación Americana de Psicología, la Asociación Americana de Investigaciones Educativas y el Consejo Nacional de Mediciones en la Educación¹⁰ cuyo objetivo es estandarizar los conceptos de evaluación, y por vez primera publicaron los “Estándares para tests y manuales para educación y psicología” (*Standards for Educational and Psychological Tests and Manuals*) con el que se aceptaron los tres términos comúnmente manejados de *Validez de contenido* (o *interna*, para algunos otros autores y libros de texto), *Validez relacionada con un criterio* (conocida también como *externa* y que incluía la *validez concurrente* y la *predictiva*), y *Validez de constructo*, (manejada por lo general con el mismo término), en virtud de los diferentes nombres utilizadas por distintos autores y estudiosos de la medición en esos años (Karmel, 1984).

Aún con ello, el manejo de distintas definiciones continuó, como la de Lemus (1977, en Carrillo 2003):

¹⁰ En inglés : American Psychological Association (APA), American Educational Research Association (AERA), & National Council on Measurement in Education (NCME)

- Se entiende por validez, el hecho de que una prueba sea de tal manera concebida, elaborada y aplicada, que mida realmente lo que se propone medir. Es la característica más importante y se refiere al valor específico del instrumento.

Posteriormente, en 1985 el comité revisó y actualizó los criterios de evaluación y publicó nuevamente los Estándares. La validez se definió entonces como “la pertinencia, significado y utilidad de las inferencias específicas derivadas de las pruebas o tests”, y la validez de un test como “el proceso de acumular evidencia para apoyar dichas inferencias” (Goodwin, 2002). Aún cuando los Estándares de 1985 seguían aceptando el uso de las tres categorías de validez establecidas en 1966, los autores de la publicación llamaron la atención sobre el punto de que el uso de estas categorías no significaba que hubiera distintos tipos de validez. De hecho, poco a poco la validez se comenzó a concebir como un concepto único tendiente a unificar las tres categorías en una sola: la de *constructo*.

Durante los 80s y 90s surgió un nuevo elemento en torno al concepto de validez el cual enfatiza la necesidad de encontrar evidencia sobre las consecuencias sociales y personales del uso de las pruebas o tests. Al respecto, Cronbach (1988, en García, 2002), apela a la importancia de la ética del evaluador y comenta “*el fondo del asunto es que los que validan tienen la obligación de revisar cuándo una práctica tiene consecuencias apropiadas para los individuos y las instituciones, y especialmente prevenir las consecuencias adversas. Usted (...) pudiera excluir la reflexión sobre las consecuencias a partir de los significados de la palabra validación, pero no puede negar la obligación de hacerlo.*”

De tal modo, encontramos definiciones ahora más amplias que intentan de alguna forma integrar el nuevo elemento, como la de Ketele (1985, en Carrillo, 2003):

- La validez de una prueba nos informa del grado en que realmente esa prueba mide el objetivo que pretende medir. Las notas que caracterizan a la validez son las siguientes:

1. Es una cuestión de grado.
2. La validez es específica para cada instrumento en particular.
3. El grado de validez de un instrumento es sensible a las características de los sujetos a los que se aplica.
4. La validez debe entenderse como un proceso continuo que requiere la acumulación de datos procedentes de distintas investigaciones y procesos.
5. Es el grado de adecuación entre lo que se quiere hacer y lo que realmente se hace.

Sin embargo, esta nueva idea de la búsqueda de evidencia de las consecuencias en la validez, provocó controversia y no fue aceptada unánimemente por los teóricos de esas décadas, por lo que el concepto no se vuelve parte de lo expuesto por estos autores en sus textos, en parte porque algunos consideraban que éste rebasaba los límites de la psicometría (Dwyer 2000 en Goodwin, 2002).

Tal es el caso de expertos en evaluación lingüística (tema principal de este trabajo), ampliamente reconocidos y citados en la enseñanza de idiomas, como Alderson, et al (1998), quienes continuaron basándose en los tres conceptos de validez aceptados en 1966 por el comité antes mencionado, y para quienes ésta tiene como propósito definir hasta qué punto una prueba mide lo que supone debe medir, y se refiere al uso o usos que se darán a las calificaciones de una prueba, o criterios de un proceso, y a las formas en las que pueden interpretarse estas calificaciones o criterios, por lo que siempre está relacionada con el propósito del objeto evaluado. Afirman que “no se puede subestimar la esencia del propósito para el que se elabora o se utiliza una prueba.” Esto es, una prueba no será válida si se utiliza con un propósito para el que

no fue diseñada. Si va a usarse con un objetivo, éste debe establecerse, y demostrarse la validez de su uso.

Sin embargo, muestran una postura abierta (y tendiente al último concepto de validez publicado en 1999 por el comité de los Estándares, como veremos más adelante) al comentar que mientras más tipos de validez se establezcan con respecto a una prueba, se puede estar más seguro de que ésta sea válida, y en su propuesta para elaborar exámenes de idiomas (1998), recomiendan, “basados en el avance en la investigación sobre la validez”, considerar y distinguir los siguientes tipos¹¹:

1. **Validez interna:** Depende del análisis lógico del contenido de la prueba para ver si contiene una muestra representativa de las destrezas lingüísticas, en el caso de un examen de lengua, que considera importantes. Las formas más comunes de evaluar la validez interna son a través de:
 - o La *validez aparente*, que se refiere a la *credibilidad* y aceptación pública de la prueba, la cual se obtiene al recoger comentarios de no expertos, quienes pueden ser los propios alumnos y administradores o profesores que no hayan colaborado en su realización. Es, por lo tanto, una opinión *intuitiva* de los usuarios.
 - o La *validez de contenido*, en la que expertos en la materia juzgan la prueba. Es “la *representatividad* o *adecuación* (sustancia, tema) de un instrumento de medida” y es la opinión de los expertos al respecto en la que se debe confiar aún cuando no coincida con la nuestra.
 - o La *validez de respuesta*. En ella se utilizan técnicas cualitativas como la introspección y la observación de sí mismos para que los candidatos puedan

¹¹ Incluimos aquí la descripción detallada sobre los tres tipos de validez manejada por Alderson, Clapham, Wall (1998) por referirse específicamente a la evaluación lingüística.

comprender y explicar cómo responden a los ítems y por qué, ya que esta información da indicaciones importantes sobre lo que evalúa la prueba (al menos para esos individuos).

2. **Validez externa.** Este tipo de validez se refiere a los estudios que comparan los resultados de los estudiantes en una prueba con la habilidad que se ha podido comprobar desde fuera. También se le conoce como *validez de criterio* ya que las calificaciones de los estudiantes se comparan con otras medidas de sus habilidades referentes a un criterio, o como *validez empírica* ya que depende de la evidencia empírica y estadística sobre si los resultados de los estudiantes en otras medidas apropiadas de sus capacidades lingüísticas son similares. De acuerdo con los autores hay dos tipos de validez externa:

- o La *validez concurrente*, que implica la comparación de los resultados de una prueba con otra medida de los mismos candidatos recogida aproximadamente al mismo tiempo de la prueba. Esta medida debe expresarse numéricamente, y puede ser una auto-evaluación de los candidatos sobre sus propias habilidades lingüísticas, o el criterio del profesor, o de otro experto.
- o La *validez predictiva* se refiere a la comparación de las calificaciones de una prueba con otra medida de los mismos candidatos pero en esta ocasión tomada un tiempo después de la aplicación de la prueba. Son comunes en las pruebas de dominio ya que intentan predecir lo bien que alguien actuará en el futuro.

3. **Validez de constructo.** Los autores de esta referencia comentan que este tipo de validez es más difícil de comprender porque tiene atributos de las dos anteriores, por lo que es considerada por algunos evaluadores como una forma superior de validez. Explican que un *constructo* se refiere a una “conceptualización teórica sobre un aspecto del

comportamiento humano que no puede medirse u observarse directamente. Ejemplos de constructo son la inteligencia, la motivación para un fin, la ansiedad, el aprovechamiento, la actitud, el dominio y la comprensión de lectura. La validación del constructo es un proceso de recogida de datos para demostrar que una prueba dada mide en realidad el constructo psicológico que sus elaboradores intentan medir. El objetivo es determinar el significado de los resultados de la prueba, para asegurar que significan lo que se cree que significan.” (Basados en Ebel y Frisbie 1991).

Ahora bien, aún cuando los debates con respecto al papel de las consecuencias sociales en la validez continúan vigentes, en la última publicación de los autores de los Estándares (AERA, APA y NCME, por sus siglas en inglés) emitida en 1999, cabe considerar que por primera vez incluyeron este elemento como forma de evidencia. Sin embargo, Goodwin y Leech (2003) destacan que no debe considerarse como una nueva forma de validez, sino como una característica que permite la validez.

De este modo, la última edición de los Estándares, describe la validez como “el grado en el que la evidencia y la teoría apoyan las interpretaciones de los resultados de las pruebas implicadas en el uso propuesto para determinada prueba.” Asimismo, indican que el proceso de validación conlleva “obtener evidencia para proporcionar una sólida base científica para interpretar los resultados propuestos para una prueba. Es la interpretación de los resultados de la prueba requeridos para determinado uso y propósito la que es evaluada, no la prueba en sí misma. Cuando los resultados se utilizan o interpretan en más de una forma, cada interpretación debe validarse.” (Goodwin y Leech, 2003).

De acuerdo con la referencia anterior, para los autores de los Estándares de 1999, las distintas fuentes de evidencia pueden esclarecer diferentes aspectos de la validez, pero no

representan distintos *tipos* de ésta, puesto que en el nuevo enfoque la validez es un concepto unitario, y se refiere a la validez de *constructo*, (i.e. la característica o concepto que una prueba o forma de medición pretende medir) dejando fuera los tres tipos anteriormente manejados como tal, aunque de una u otra forma se conservan dentro de los cinco tipos de *evidencia de validez* ahora definidos y que a continuación se resumen:

1. Evidencia basada en el contenido.

Este tipo de evidencia se basa en el análisis de las evaluaciones de expertos en cuanto al contenido del instrumento de medida, e incluye ítems, tareas, formatos, términos y el proceso al que se somete a los evaluados. La revisión de expertos pretende obtener evidencia de las características de un test como suficiencia, claridad, relevancia, así como la correspondencia entre las tareas y los reactivos y la definición del constructo.

2. Evidencia basada en los procesos de respuesta.

Examina el grado en que las tareas o tipos de respuesta requerida por los evaluados corresponde al constructo establecido o buscado. Para hacerlo se realizan observaciones de los examinados conforme realizan las tareas y se les entrevista para determinar las razones de dar ciertas respuestas. Debe procurarse que los observadores, jueces o examinadores utilicen los criterios establecidos y no incluyan factores externos e irrelevantes que no concuerden con las interpretaciones esperadas de los resultados.

3. Evidencia basada en la estructura interna.

Evalúa el grado en que los componentes internos de la prueba concuerdan con el constructo establecido, y por lo general se estima a través del análisis factorial, a través del cual se pueden identificar los factores que tienen en común los distintos grupos de

ítems. Los evaluadores deben tener cuidado de no basarse únicamente en el resultado del análisis factorial para confirmar si existe o no evidencia de validez.

4. Evidencia basada en la relación con otras variables.

Esta forma de evidencia engloba algunas de los tipos de validez aceptados en los Estándares de 1985, como la validez relacionada con un criterio (que, como se mencionó previamente, incluía la validez concurrente y la predictiva), y la validez de constructo (conformada por la validez convergente y la discriminante). Para obtener este tipo de evidencia se utilizan comúnmente estudios correlacionales (sobre la naturaleza y grado de la relación entre los resultados y variables externas); estudios comparativos de grupos con un criterio establecido (para probar hipótesis sobre diferencias esperadas entre grupos con diferentes condiciones); y estudios experimentales (para comprobar hipótesis sobre los efectos de intervenir en los resultados, así como la eficacia en las decisiones sobre colocación, selección y clasificación).

5. Evidencia basada en las consecuencias de la evaluación.

Este tipo de evidencia de validez pertenece a las consecuencias anticipadas e imprevistas, tanto positivas como negativas, de la medida. Es decir, se aplica una prueba con el fin de beneficiarse, de alguna manera, con los resultados (por ejemplo para elegir a un candidato idóneo para un trabajo, o para prevenir que alguien con un perfil incorrecto sea seleccionado para ello, etc).

Como puede observarse, el término de *validez* es sumamente amplio y está conformado por muchos y diversos elementos que a lo largo de los años han generado su transformación. A pesar de la nueva visión establecida en los Estándares de 1999, la mayoría de los expertos,

investigadores, libros de texto y académicos, continúan manejando el de los tres tipos de validez y, como señala Goodwin (2003), el cambio sólo podrá ser gradual.

- Relación entre confiabilidad y validez

La relación entre estos conceptos puede ser recíproca o no, por lo que puede pasar de ser algo simple a algo complejo. Hay que mencionar que si una prueba o proceso no es válido, no puede ser confiable, pues si la prueba no mide el objetivo que pretende medir en forma consistente, no lo puede medir en forma precisa. Sin embargo, una prueba sí puede ser confiable aunque no sea válida. Es decir, puede arrojar los mismos resultados de manera consistente aunque no mida lo que se supone debe medir. Por lo tanto, aunque se necesita confiabilidad para obtener validez, no se requiere validez para la confiabilidad, aunque la confiabilidad por si misma no es suficiente.

Alderson, et al (1998) indican que en la práctica, “ni la validez ni la confiabilidad son valores absolutos: hay grados de ambas y es habitual hablar de un compromiso entre las dos”. Comentan que en ocasiones puede decidirse maximizar una en detrimento de la otra y que la que se decida maximizar dependerá del objetivo de la prueba y de las consecuencias que esto pueda traer a los candidatos. Concluyen diciendo que si un proceso en particular discierne entre validez o fiabilidad no importa, que lo importante es que seamos conscientes de los problemas que, como evaluadores, se nos presentan: “Puesto que una prueba no puede ser válida sin ser fiable, es esencial que las pruebas se analicen todo lo que se pueda para garantizar su confiabilidad, y si algunas de las formas de análisis resulta que comprueban la validez y no la confiabilidad, al menos nos dan más información sobre la prueba. Lo que importa en realidad es si la prueba da un resultado que puede considerarse como un reflejo justo y preciso de la habilidad lingüística del candidato.”

▪ Evaluación del aprendizaje de lenguas

Para comenzar diremos que, sobre todo hoy en día, los exámenes lingüísticos juegan un importante papel en la vida de las personas al ser una puerta hacia oportunidades educativas y de empleo, así como a experiencias personales como el turismo o el vivir en distintos países.

Recordemos que la evaluación del lenguaje es una forma de obtener evidencia sobre la actuación, competencia, conocimiento y habilidad que una persona hace de una lengua, ya sea extranjera, materna, o segunda (y hasta tercera) lengua.

No todos los exámenes de lengua son iguales. Como veremos más adelante, difieren con respecto a su diseño (o tipo), y por lo que se refiere a su objetivo (propósito). Sin embargo, cualquiera que sea el método o diseño, la evaluación del lenguaje se basa en teorías sobre la naturaleza de su uso en determinado contexto, lo que se verá reflejado a lo largo de la prueba.

Es importante mencionar que las teorías del lenguaje y de su uso han cambiado a lo largo de los años y que debido a ello los exámenes o pruebas tenderán a diferentes orientaciones teóricas con respecto al significado de *la lengua*, a lo que se entiende por *aptitud lingüística*, a lo que significa *aprender* o *adquirir* una lengua, y a lo que hacen los hablantes nativos con ella. (Alderson et al 1998)

El uso del lenguaje puede definirse como “la creación o interpretación de significados específicos en el discurso de una persona, o como la negociación interactiva y dinámica de cierto significado entre dos o más individuos en una situación particular” (Bachman y Palmer, 1997). Además, existen diferentes perspectivas desde las cuales se debe considerar el uso del lenguaje. Una de ellas es la perspectiva lingüística, que considera que el lenguaje está formado por sistemas como la gramática, los tiempos, la semántica, la fonología, la fonética, la entonación, el discurso, etc., y que se relacionan entre sí y con otros dominios no lingüísticos. La segunda es

desde la psicolingüística, que considera al lenguaje como un proceso cognitivo asociado con el desarrollo, el aprendizaje y el pensamiento y con los procesos de producción. Por último, encontramos la consideración sociolingüística, que asume al lenguaje como inherente al tiempo y el espacio para funcionar dentro de los contextos y procesos sociales (Davies, 1990).

Para Lado (1964), la teoría de la evaluación del lenguaje percibe al lenguaje como un sistema de hábitos comunicativos. Estos hábitos incluyen cuestiones de forma, significado y distribución en distintos niveles de estructura, fundamentalmente en lo que se refiere a la oración, las cláusulas, frases, palabras, morfemas y fonemas. La teoría asume que al evaluar el control de los problemas se está evaluando el control (conocimiento) del lenguaje.

Cada teoría contiene como principales componentes ciertos *constructos* (o conceptos psicológicos) que se reflejan en la evaluación y para los que deben encontrarse evidencias de validez.

La taxonomía que Bachman propone en *Fundamental Considerations in Language Testing* (1990) es un marco de referencia teórico que permite analizar distintas pruebas, y que ha servido de base para la evaluación de exámenes de lengua (Alderson, et al, 1998). En él, separa la habilidad lingüística comunicativa en dos clases de competencias:

1. Competencia Organizativa, que incluye:
 - a. la competencia gramatical (vocabulario, morfología, sintaxis, fonología/grafología).
 - b. la competencia textual (cohesión y organización retórica).
2. Competencia pragmática, que se refiere a:
 - a. la competencia ilocutiva (el uso de funciones)

b. la competencia sociolingüística (la sensibilidad hacia diferencias de dialecto o variedad; hacia las diferencias de registro; hacia la naturalidad; y la habilidad para interpretar referencias culturales y lenguaje figurado).

En virtud de que, por su propia naturaleza, el lenguaje se presenta en situaciones muy diversas y particulares, cada acto de habla es virtualmente único, y sería imposible enlistar todos los posibles ejemplos de ellos. Sin embargo es factible identificar ciertas características distintivas del uso del lenguaje y utilizarlas para describir sus dominios. Por lo tanto se deben realizar inferencias sobre las habilidades del lenguaje de quienes van a ser evaluados con el fin de generalizarlas hacia los dominios específicos en que los candidatos puedan requerir utilizarlo.

De este modo, en los exámenes de lengua se debe seleccionar una muestra de lo que se considera la habilidad lingüística que se debe conocer o haber aprendido, tomando en cuenta que prever las dificultades que pueden surgir es conocer el lenguaje. Dicha muestra deberá ser elegida al azar de un repertorio inmenso de posibilidades.

- Objetivos de los exámenes de lengua.

Cuando evaluamos, lo hacemos por distintas razones, por lo que los exámenes de lengua pueden tener diversos propósitos, aunque el fin último siempre sea lograr una evaluación justa y pertinente. De acuerdo con su objetivo, se clasifican como exámenes:

- de progreso (*progress tests*). Este tipo de pruebas funcionan como las formativas ya que se aplican a lo largo del curso para monitorear el avance del alumno, así como para detectar si tiene problema con algún tema y en tal caso apoyarlo con nuevas explicaciones, ejemplos o prácticas adicionales.
- de logro o aprovechamiento (*achievement tests*). Podríamos afirmar que son como los *sumativos*, ya que pretenden establecer lo que los alumnos han aprendido en

determinado curso, por lo que se aplican al final y por lo general comprenden todos o la mayoría de los temas del programa de estudios.

- de dominio (*proficiency tests*). Estos exámenes, por el contrario, no se basan en ningún programa específico, ya que están diseñados para evaluar las distintas competencias lingüísticas de los evaluados, sin importar dónde o cómo las hayan adquirido. En este tipo de pruebas se toma en consideración las especificaciones de lo que un candidato debe poder hacer para considerar que tiene dominio de la lengua. Tal es el caso de exámenes como el TOEFL test, Cambridge First Certificate, Cambridge Proficiency, etc.
- de diagnóstico (*diagnostic tests*). Estas pruebas pretenden identificar las áreas en que un alumno necesita ayuda complementaria, por ejemplo en alguna de las cuatro habilidades: leer, escribir, escuchar o hablar. También pueden ser más precisas si se trata de determinar las áreas gramaticales específicas en las que el candidato tenga alguna debilidad. De acuerdo con Alderson, et al (1998), es muy difícil realizar una prueba de este tipo debido a las complejidades de la competencia lingüística, razón por la cual “existen muy pocas pruebas de diagnóstico *puras*. Sin embargo se usan con frecuencia pruebas de aprovechamiento y de dominio, aunque de forma no sistemática, con el propósito de una prueba de diagnóstico”.
- de colocación (*placement tests*). Los exámenes de colocación o de nivel son definidos por Harrison (1983 en Murray, 2001) con el objeto de “colocar a los nuevos estudiantes en grupos de enseñanza para que puedan iniciar un curso aproximadamente en el mismo nivel de los otros estudiante de la clase [...] y están relacionados con el estado general de la habilidad que tenga de la lengua el candidato en el momento de la evaluación, y no con un tema de aprendizaje específico.”

En virtud de que en este trabajo nos ocupa analizar el proceso de clasificación y el examen de colocación escrito que se utiliza en la UIA, más adelante hablaremos con más detalle sobre las evaluaciones con este fin.

- Tipos de exámenes

Existen ciertas categorías que permiten distinguir entre distintos tipos de exámenes de lengua (Henning, 1987)¹², como son:

- Exámenes Objetivos Vs. Subjetivos

Se distinguen por su forma de calificación. Los primeros utilizan reactivos de respuesta estructurada (opción múltiple, verdadero-falso, reconocimiento del error, etc.), en los que el candidato debe elegir una respuesta entre varias posibles, que será considerada como correcta o incorrecta sin posibilidad de términos medios o de que exista otra alternativa. Por lo general se cuenta con una *clave de respuestas* que permite su rápida revisión.

En el caso de las pruebas *Subjetivas*, el tipo de reactivos o tareas no permiten que los examinadores apliquen juicios del tipo anterior, es decir *correcto-incorrecto*. Deben evaluar la calidad de la respuesta de un candidato a una tarea (como una composición, una entrevista, *cloze tests*, en los que se deben completar espacios que permiten diferentes respuestas, etc.). Para calificar este tipo de pruebas, los evaluadores deben, idealmente, utilizar una *escala de valoración* en la cual basarse para evitar emitir juicios personales o considerar aspectos ajenos a la evaluación. Dicha escala puede consistir en números, letras o etiquetas que indiquen el tipo de actuación lingüística al que se refiere cada punto de la escala (por ejemplo: Excelente: *inglés natural con errores mínimos*;

¹² Nos basamos en este autor por considerar que la comparación que hace entre categorías opuestas, permite un mejor entendimiento de las mismas.

Bien: *realización simple aunque precisa de la tarea encomendada, con algunos errores serios, etc.*).

- Exámenes Directos vs. Indirectos

Este tipo de examen se diferencia en cuanto a su grado de naturalidad comunicativa. Por ejemplo una entrevista sería una evaluación mucho más directa, que una prueba en la que se completen oraciones o se elijan alternativas puesto que, al no tener un contexto, hacen uso del idioma de forma menos natural en cuanto a la habilidad de comunicación.

- Exámenes de Puntos Discretos vs. Evaluación Integradora

Los primeros miden el conocimiento o desempeño en áreas restringidas de la lengua meta. Es decir, cada pregunta de la prueba está diseñada para medir un punto de contenido diferente a los demás. Este tipo de exámenes funcionan adecuadamente como diagnóstico. En el modelo de evaluación integrador se utiliza de manera simultánea una mayor variedad de habilidades del lenguaje. Estas pruebas evalúan de manera más general el dominio del candidato de la lengua. Ejemplo de los primeros serían completar oraciones para evaluar el uso de algún tiempo verbal. De los segundos, una entrevista, un dictado, escribir un ensayo, etc.

- Exámenes referidos al Criterio vs. referidos a la Norma

Los que se refieren a un Criterio son concebidos antes de la instrucción. La prueba debe concordar con los objetivos de enseñanza, por lo que se permite que el profesor enseñe con base en el examen para asegurarse de que los objetivos sean alcanzados. Asimismo, los criterios de aprobación se establecen por adelantado y quienes no cumplen con ellos requieren repetir el curso. El logro del estudiante se mide con respecto al grado de su

aprendizaje o manejo del contenido especificado previamente. No se compara a los alumnos entre sí, sino que se les compara con un nivel de aprovechamiento.

Los exámenes referidos a la Norma en principio deben haber sido previamente aplicados a una muestra amplia de la población a evaluar. Los estándares de aprovechamiento se determinan después de que el examen ha sido diseñado y aplicado. Dichos criterios se basan en el resultado promedio de otros estudiantes de la misma población. En virtud de que se espera contar con un amplio rango de distribución de resultados, se incluyen reactivos de distintos niveles de dificultad con el fin de poder distinguir entre los de alto y bajo desempeño. Tienen como objetivo establecer el orden que ocupan los candidatos en una escala ordenada. De este modo el significado de la puntuación de un individuo adquiere sentido al considerar el nivel común de ejecución de un grupo. Por lo general en las pruebas que se aplican con respecto a la norma los puntajes pueden sumarse para dar un resultado total y así ordenar a los alumnos de acuerdo con ello.

- Exámenes de Velocidad vs. Exámenes de Poder

Los primeros son considerados hasta cierto punto fáciles, por lo que se espera que cualquier persona logre tener todas las respuestas correctas en un tiempo límite. Se compara a los candidatos de acuerdo con su rapidez y no únicamente con respecto a su conocimiento. Los segundos utilizan reactivos de mayor dificultad pero se da el tiempo suficiente para que todos los candidatos terminen, sin embargo se asume que no todos logran tener todas las respuestas correctas.

▪ **Consideraciones adicionales**

Con base en el tema que nos ocupa en este trabajo encontramos pertinente ampliar algunos aspectos, los cuales se refieren al examen de colocación y a la importancia del proceso para llevarlo a cabo, a los exámenes de opción múltiple, y a la problemática de la evaluación subjetiva de la prueba, es decir la entrevista en la LM.

- El examen y el proceso de colocación

Comentaremos cuatro aspectos: la situación que priva en cuanto a la elaboración de los *exámenes de colocación*; algunas definiciones; los componentes lingüísticos que tienden a integrar estas pruebas; y los elementos para validarlas.

1. Para empezar debemos mencionar que en muchas instituciones (y muy comúnmente en el contexto Mexicano), la elaboración de los exámenes de colocación y el proceso que conlleva su aplicación son realizados de manera irregular y sin ninguna base u orden. En ocasiones sucede que los encargados de elaborar la prueba cuentan con poca o ninguna experiencia y conocimiento sobre evaluación para poder llevarlo a cabo de manera eficiente (Murray, 2001). De hecho, con frecuencia algunos profesores de lengua no están familiarizados con el proceso que implica la colocación de los estudiantes en los cursos (Kim y Shin, 2006) por lo que les puede ser difícil distinguir entre la perspectiva a utilizar para evaluar el desempeño y progreso de los estudiantes durante un curso, y el tipo de enfoque requerido para evaluar a un candidato que va a ingresar al mismo.

Estos son factores que tienen como consecuencia que el grado de validez, confiabilidad, y precisión de este tipo de instrumentos (y en ocasiones del procedimiento completo de clasificación) llegue a ser cuestionable.

La realidad es que, aún cuando el examen de colocación es una de las pruebas de lengua más comúnmente utilizadas dentro de las distintas instituciones dedicadas a la enseñanza de idiomas, poco se ha escrito específicamente sobre la elaboración de este tipo de pruebas que pueda ayudar a instituciones y profesores a realizarlos (Fulcher, 1997; Murray 2001, Green y Weir, 2004). De hecho, aunque existe una vasta literatura sobre evaluación de idiomas, estos textos tratan de manera muy general el tema sobre la creación y evaluación de este tipo de exámenes. De igual modo, la naturaleza y validez de los exámenes de colocación no son comúnmente comentados en la literatura disponible probablemente debido a que son pocas las investigaciones al respecto, y las que se realizan se hacen a nivel local dentro de las instituciones y por lo general no llegan a publicarse (Fulcher, 2006¹³).

Sin embargo, contamos con los principios de la evaluación, con las teorías de autores ampliamente reconocidos en los terrenos de la teoría lingüística y de la evaluación de lenguas, así como con paquetería informática (como el *SPSS* o *ITEMAN*, por ejemplo), para realizar nuestras propias investigaciones al respecto y mejorar nuestro desempeño para la elaboración de exámenes de colocación y para establecer el procedimiento más adecuado para aplicarlos adecuadamente.

2. Entre las definiciones de *examen de colocación* encontramos que para Brown (1989) es un elemento importante para colocar a los estudiantes en grupos de habilidades lingüísticas relativamente homogéneos. Chandavimol (1988, en Murray 2001) establece dos propósitos fundamentales de la clasificación: el primero asignar a los alumnos a clases con diversos niveles de dificultad –desde principiante a avanzado; y el segundo, determinar las necesidades de los estudiantes para así proporcionarles un mayor grado de flexibilidad para cubrir dichas necesidades. Agrega que los exámenes de colocación determinan si un

¹³ Comunicación personal vía correo electrónico con el Dr. Glen Fulcher, (Presidente de ILTA *International Language Testing Association*; reconocido ponente e investigador del *Center for Language Teacher Education and Applied Linguistics, School of Education at the University of Leicester*), en la que indica que no se ha hecho mucho del año 2000 a la fecha sobre exámenes de colocación (6 de octubre, 2006).

estudiante domina ciertas habilidades para realizar el programa de un curso, y que permiten identificar debilidades o problemas en su uso de la lengua con el fin de enfocarse y trabajar en ellas.

Para Alderson, et al (1998), los exámenes de colocación de los alumnos que ingresan a un programa de estudios de lengua tiene como objetivo evaluar los niveles de habilidad lingüística con el fin de colocarlos en el curso o clase apropiada.

Green y Weir (2004) señalan que los exámenes de colocación juegan dos papeles principales: ayudar a la creación de grupos de estudiantes con habilidades lingüísticas homogéneas, y comprobar el dominio de un alumno en cuanto al contenido de un curso.

Brown (2005, en Kim y Shin 2006) enfatiza que un buen examen de colocación debe separar, con un alto nivel de confiabilidad y validez, a los estudiantes en niveles apropiados.

Podemos resumir que el objetivo esencial de los exámenes de colocación de idiomas es clasificar o situar a los candidatos que lo aplican dentro de un continuo de categorías correspondientes a requisitos específicos en el manejo de la lengua y relacionados con los niveles de aprendizaje que establezca un programa de estudios. Y hay que hacerlo con el menor margen de error posible.

3. Con respecto a los componentes lingüísticos encontrados en este tipo de pruebas encontramos que pueden ser muy distintos, puesto que no existen criterios determinados en cuanto a qué se deba evaluar en ellos.

Alderson, et al (1998) indican que estos exámenes pueden estar basados en aspectos del programa de enseñanza establecido por la institución, o en otro tipo de materiales no relacionados con éste.

Por su parte Murray (2001) hace referencia a los distintos componentes que pueden encontrarse en exámenes de colocación, como son:

- la comprensión auditiva, que mide la habilidad de utilizar el propio conocimiento del lenguaje para entender textos orales.
- la producción oral, a través de la cual se mide la habilidad comunicativa del candidato.
- la gramática, la cual es incluida en muchos exámenes de colocación ya que este constructo permite medir un aspecto importante del desempeño general de un individuo en cuanto a su manejo del lenguaje.
- el vocabulario, cuya evaluación busca por lo general medir el conocimiento de las palabras en cuanto a sus diversos significados, o a medir el tamaño del léxico del candidato con el fin de saber si puede manejar el lenguaje de una clase en particular. Medir ese componente puede ser útil para cursos con propósitos específicos en algún área en particular, como arquitectura, ingeniería, medicina, aeronáutica, etc.
- la comprensión de lectura, que pretende medir la habilidad de un lector para extraer cierto grado de significado bajo condiciones de desempeño específicos. Este componente es comúnmente utilizado en exámenes de colocación, aunque por sí mismo no permite evaluar el grado de producción (ya sea escrita, u oral). La posible problemática que Murray resalta es la de encontrar textos auténticos (o semi-auténticos) que permitan evaluar el desempeño del candidato en contextos reales.
- la producción escrita. A través de este componente se pretende evaluar el grado en el que el individuo puede hacer uso efectivo de los distintos y complejos aspectos del uso del lenguaje de manera clara, por lo general a través de ensayos. Una posible desventaja

Entre varios ejemplos de exámenes de colocación y los diversos componentes en sus pruebas, reportamos los siguientes:

- El examen aplicado en *The English Language Institute* (ELI) de la universidad de Surrey tiene como propósito seleccionar a aquellos estudiantes que requieren entrevista oral con el fin de que sean los menos, considerando que en un año escolar pueden llegar a tener un promedio de 1600 alumnos o más. Incluye: un ensayo descriptivo (sin tema obligatorio); un ensayo argumentativo (a elegir de tres opciones); una prueba de estructura gramatical de diez reactivos, una sección de lectura de comprensión de ocho. (Fulcher, 1997).
- En el caso del examen de VELI (*Vancouver English Language Institute*), Murray (1999) sugiere, después de su análisis del examen original que incluía vocabulario, gramática y comprensión de lectura, incluir *subtests* que evalúen puntos gramaticales, lectura, comprensión auditiva y escritura, cada uno con 50 reactivos (con base en Brown, 1995), además de incluir una entrevista oral.
- El *Advanced Placement International English Language* (APIEL) pretende medir el dominio de hablantes no nativos de inglés para evaluar su posible participación en una comunidad académica de habla inglesa.

Los componentes que evalúa son comprensión auditiva y de lectura; producción oral y escrita. El formato incluye una sección de opción múltiple para los primeros dos componentes, y una segunda sección de respuesta libre para escribir dos ensayos en

ochenta minutos. La sección oral incluye cinco preguntas y se graban las respuestas de los candidatos.

Entre las desventajas reportadas encontramos que se aplica una vez al año, tiene una duración efectiva de tres horas (más el tiempo de los trámites administrativos), y los resultados se entregan cinco meses después, por lo que no es práctico para instituciones que inician programas con mayor frecuencia al año. Otros problemas son la falta de naturalidad en los pasajes de comprensión auditiva, en los que se hace evidente que los textos fueron leídos. (Cho, 2002).

- El *Quick Placement Test* de *Oxford University Press*. Está diseñado para ayudar a profesores y administradores de cursos a tomar decisiones rápidas sobre el nivel en el que los candidatos deben ser colocados, o sobre si pueden ingresar a cursos específicos. Cuenta con dos versiones, sobre papel y por computadora, ambas de opción múltiple. La prueba sobre papel toma alrededor de treinta minutos en ser aplicada y se evalúan la comprensión oral y de lectura, así como estructuras, incluyendo gramática y vocabulario. Un posible problema de la prueba es que cada institución debe desarrollar sus propias escalas para interpretar los resultados de manera significativa para la institución y así determinar el nivel correspondiente (Ardeshir, 2003), lo que deja dudas con respecto a su exactitud.
- El GPT (*The Global Placement Test*), un examen sobre papel utilizado en diversas escuelas a nivel mundial, mide exclusivamente estructuras gramaticales, y está diseñado para reflejar las necesidades de los estudiantes con respecto a los contenidos de los cursos. Sobre la prueba se cuestiona si la secuencia de los elementos de la competencia gramatical que presenta debería seguir el orden que aparece en los textos de enseñanza. (Green y Weir, 2004).

- En el examen en línea de la universidad de Stanford, los elementos que se evalúan son gramática y vocabulario (a través de opción múltiple o llenado de espacios). Posteriormente los candidatos deben escribir su autobiografía, y leer un párrafo corto con preguntas de opción múltiple o escribir un resumen del mismo. La sección en línea es calificada antes de la llegada de los estudiantes a la universidad. Una vez ahí y con el resultado previo de cada prueba, deben pasar a la entrevista oral para determinar finalmente su ubicación. (Bernhardt, et al, 2004)
- El examen de colocación de la Universidad de Columbia desarrollado en el *Teachers College* conocido como *CEP (Community English Program placement test battery)* evalúa las secciones de gramática, lectura, comprensión auditiva, producción oral y escrita. (Park, 2004).

Como comentamos previamente, no existen criterios absolutos para decidir cuáles componentes incluir en un examen de colocación, pero debemos resaltar que una evaluación de este tipo debe mantener criterios de practicidad, puesto que pretender evaluarlos todos puede implicar un gran consumo de tiempo.

4. Con respecto al último punto de esta sección, nos referiremos a la validación de este tipo de pruebas. En un estudio empírico publicado al respecto por Wall, et al (1994) se plantea que para evaluar y validar un examen de colocación debemos preguntarnos lo siguiente:

- ¿El examen de colocación identifica correctamente a los estudiantes que necesitan el estudio del idioma?
- Los candidatos que realizan la prueba sienten que su nivel de idioma ha sido apropiadamente evaluado? (Validez aparente).

- ¿El contenido de la prueba es apropiado para lo que ésta se utiliza? (Validez de contenido).
- ¿Es un examen confiable?

Sin embargo, y como establecimos en el planteamiento del problema, desde nuestro punto de vista el examen de colocación no es un elemento único y definitivo para determinar en dónde ubicar al estudiante. Debemos considerar que la aplicación de una prueba de esta índole se realiza dentro de un proceso que inicia con los aspectos administrativos de la misma, pasando por la calidad y presentación del o los instrumentos de evaluación (prueba escrita, oral, de comprensión auditiva, de lectura, etc., en cualquiera de sus tipos), el lugar y ambiente en el que se realiza, el tiempo y hora del día para hacerlo, y la manera de calificar e interpretar resultados por parte de los evaluadores, y que, aunque puede haber quien lo olvide, incluye a los mismos evaluados, pues como Bradshaw (1990) comenta: las reacciones de éstos hacia los procesos de evaluación pueden tener efectos en los resultados.

Por lo anterior, como pregunta integradora a los planteamientos de Wall, et al (1994) mencionados arriba, podemos retomar la esencia de la pregunta de investigación establecida para el presente trabajo, y agregar:

- ¿Cuál es la efectividad en cuanto a evidencia de validez y grado de confiabilidad del proceso de colocación para establecer el nivel de conocimiento de los estudiantes y poder ubicarlos en el programa de estudios adecuado?

- **Los exámenes de opción múltiple**

Profundizaremos en este tipo de formato puesto que es el que utiliza el examen analizado en este trabajo.

Los reactivos de opción múltiple son frecuentemente utilizados en la elaboración de pruebas objetivas. Evidentemente existen ventajas, desventajas y controversia asociadas con los distintos tipos de instrumentos de evaluación, y el uso de la opción múltiple no está exento a ello. Lo importante es tomar en cuenta que no hay un formato establecido para medir todo lo que se puede evaluar y que debemos decidir qué tipo de prueba utilizar dependiendo de nuestro objetivo.

Una de las ventajas es que nos permiten realizar un gran número de preguntas para obtener un buen muestreo del conocimiento de los candidatos en un lapso corto de tiempo. Otra ventaja muy apreciada por los evaluadores es, sin duda, su fácil y rápida corrección. Una más se relaciona con su alto grado de confiabilidad (siempre y cuando los reactivos estén bien escritos, por supuesto). De hecho algunos expertos consideran que un buen reactivo permite establecer un contexto realista, natural y confiable para evaluar el conocimiento de los estudiantes (Ebel y Frisbie (1991).

Una de las críticas al uso de la opción múltiple es que los estudiantes pueden llegar a la respuesta correcta a través de la eliminación. Sin embargo, Ebel y Frisbie (1991) afirman que la mayoría de los especialistas en evaluación del constructo aseguran que el conocimiento y la habilidad que se requiere para eliminar alternativas incorrectas es el mismo que se utiliza para elegir la respuesta adecuada, y que el hecho de que un estudiante responda de manera analítica y haga uso de su capacidad de pensamiento crítico y de resolución de problemas, habla de un proceso racional que debe ser aplaudido, más no criticado. Nos dicen que se ha comprobado que una persona que no tiene idea de la respuesta correcta no tendría el conocimiento suficiente para eliminar con seguridad las alternativas erróneas. Concluyen que el hecho de elegir la respuesta correcta por eliminación, no constituye un problema serio, pero resaltan que esto será así siempre y cuando el reactivo esté bien construido.

De hecho, la creencia errónea más común sobre este tipo de exámenes es que son fáciles de formular. La tendencia de quienes no tienen la experiencia para hacerlo es utilizar distractores que no tienen ninguna relación con la respuesta correcta, por lo que en esos casos es fácil para los estudiantes descartar las alternativas *disparadas*.

A continuación resumimos algunas recomendaciones para escribir buenos reactivos (con base en Alderson et al, 1998, y McDonald, 2001):

- La *base* (o *tallo*) del ítem es la parte fundamental del reactivo. La *base* debe ser clara, concisa, centrada en un objetivo y debe tener un enfoque positivo. Si las personas evaluadas no entienden la base, no podrán responder la pregunta.
- La respuesta *correcta* debe ser solamente una, y debemos asegurarnos de que en verdad lo sea. Esto quiere decir que deben evitarse alternativas ambiguas que permitan otras posibles respuestas. Para evitar estos problemas recomiendan que otros expertos revisen los reactivos y sus alternativas y confirmen que no hay posibilidad de controversia.
- Cada opción debe encajar perfectamente en la base del reactivo con el fin de que no sea evidente cuál es correcta o cuáles erróneas.
- Cada alternativa errónea debe tener posibilidades de elección por parte de los examinados (se recomiendan por lo menos cuatro para que las posibilidades de elección al azar sea sólo del 25%. Por supuesto, un reactivo con mayor número de opciones puede discriminar mejor, pero es difícil encontrar cinco o más alternativas que funcionen como se espera).
- Si nadie elige alguna de las opciones significa que ésta es inútil y que debe eliminarse. Recordemos que el objetivo de los distractores es discriminar entre los individuos que saben menos y los que saben más, razón por la cual es de gran importancia que las

opciones sean de alguna forma atractivas para quienes poseen un menor dominio del tema evaluado. Si la mayoría de quienes poseen un mayor conocimiento del idioma eligen alguna alternativa incorrecta, o por el contrario, si los de bajo nivel eligen la respuesta correcta, significa que el reactivo no está funcionando adecuadamente.

- De manera ideal, las opciones deben ser parecidas y preferentemente del mismo largo. Opciones más largas que otras tienden a ser elegidas o eliminadas por los estudiantes porque tienden a creer que es una especie de ‘clave’ para elegirla o no.
- Se deben evitar opciones negativas porque tienden a incrementar la dificultad no con respecto al conocimiento del estudiante, sino porque son engañosas. De hecho no se consideran distractores aceptables.

Consideramos relevante tomar en cuenta estos factores para elaborar o evaluar exámenes de opción múltiple pues es un hecho que los reactivos que no cumplan con los objetivos de la prueba definitivamente afectarán su validez. De igual modo si se incluyen ítems que produzcan ambigüedad se evitará que los examinados descifren lo que se pretende preguntar y la confiabilidad del examen se verá disminuida.

- La evaluación oral y su problemática

Consideramos relevante ampliar este tema puesto que el examen de colocación que nos ocupa está compuesto de una prueba escrita y una entrevista en LM, como veremos con más detalle en el Marco Contextual en el siguiente capítulo.

En principio es importante recordar que un aspecto relevante para lograr comunicarnos en una lengua extranjera o segunda lengua tiene que ver con el dominio de un sistema de sonidos diferente al de nuestro idioma nativo. Para ello, debemos controlar el ritmo, la entonación y los patrones de acentuación para que nuestros interlocutores puedan entender nuestro mensaje.

Aunque no sucede en todos los idiomas, en algunos como el inglés también es importante pronunciar de manera precisa los sonidos consonantes y vocales puesto que de no ser así, podemos crear interferencias en cuanto a la comprensión de nuestro mensaje. Una vez dominados estos aspectos, es posible hablar una LE o L2 con un alto grado de desempeño lingüístico a pesar de mantener el acento propio de la lengua nativa, lo que no constituye un impedimento para hacerlo.

Ahora bien, cuando se evalúa a los estudiantes con respecto a su habilidad para comunicarse oralmente, se pone especial atención en su capacidad para transmitir su mensaje o intención comunicativa. Lo más importante a considerar es si la persona logra hacerse entender. Los exámenes orales por lo general le indican al estudiante sobre qué hablar. Sin embargo su elección de palabras y estructuras depende totalmente de él, de su personalidad (si es extrovertido o tímido) y esencialmente de su competencia lingüística.

Como indica Fulcher (1997), quien se dedica a la evaluación sabe que todos los exámenes contienen errores, y es por esto que, en el caso de los exámenes de colocación, la entrevista juega un papel fundamental para confirmar que la prueba escrita haya ubicado apropiadamente a los candidatos. Por lo tanto, a pesar de ser un elemento de confirmación, debemos tomar en cuenta que en pruebas subjetivas como ésta existen factores que pueden influir en su grado de validez y confiabilidad, como son: el evaluado, su interlocutor (i.e. el examinador), los criterios de evaluación, la tarea asignada, y la interacción entre estos elementos (O'Sullivan et al, 2002).

No obstante, el mito sobre la entrevista es que se trata de una forma rápida y fácil de evaluar la competencia lingüística de una persona. Lo que normalmente sucede en estos casos es que el diálogo gira en torno de temas superfluos que sólo requieren de un vocabulario limitado y típico, así como del uso de estructuras sin alto grado de complejidad, lo que permite que algunos candidatos incluso anticipen y practiquen previamente sus respuestas. La realidad es que los

exámenes orales que no se planean adecuadamente por lo general son instrumentos poco confiables para evaluar el desempeño lingüístico de los candidatos.

Como indican Alderson, et al (1998), la entrevista debe cubrir de manera cuidadosa aquellos aspectos de la prueba considerados importantes para evaluar a cada candidato de forma similar. En muchas ocasiones quienes aplican esta sección del examen olvidan o no toman en cuenta los factores de ética e imparcialidad requeridos para evaluar con los mismos criterios a todos los candidatos. Es decir, suele ocurrir que a algunos de ellos se les pide hacer comentarios apropiados pero simples, mientras que de otros evaluados se espera que utilicen un lenguaje mucho más complejo, lo que puede hacerlos incurrir en errores que el evaluador toma en cuenta para establecer su decisión final sobre su habilidad oral.

De acuerdo con el estudio realizado por Brown (2003), los entrevistadores tienen mucho que ver con el desempeño de un candidato, ya que suelen cambiar su manera de estructurar preguntas o temas, así como sus técnicas para cuestionarlos e incluso su manera de retroalimentarlos. Afirma que las diferencias que existen en su forma de actuar, (es decir, la falta de unificación de criterios) tienen como resultado distintas impresiones sobre la habilidad del evaluado, así que el mismo individuo puede obtener un resultado distinto en cuanto a su competencia comunicativa si es evaluado por dos o más entrevistadores.

De igual modo, Kim y Shin (2006) afirma que el comportamiento de los evaluadores es uno de los factores más importantes para la variación de resultados en los exámenes orales y, por ende, en la validez y confiabilidad de éstos. Comentan que la diferencia de resultados se presenta no sólo de entrevistador en entrevistador, sino también en cuanto a que en ocasiones un mismo evaluador no es consistente al utilizar sus propios criterios.

Ahora bien, en teoría, un buen entrevistador debe estar consciente de la importancia de ser imparcial con todos los candidatos y de evaluarlos con los mismos criterios. Esto forma parte de

la ética en la evaluación y del concepto de justicia que debe existir en torno a ella. Desafortunadamente en la práctica, en ocasiones un entrevistador no alcanza si quiera a darse cuenta de que puede estar actuando imparcialmente, y esto puede deberse simplemente a la falta de capacitación para llevar a cabo este tipo de evaluación.

Entre otras cosas, quien evalúa la habilidad oral dentro del proceso de colocación debe lograr que los candidatos se sientan relajados para que puedan dar ejemplo de una comunicación auténtica, lo que puede alcanzarse al mostrar su interés en cada entrevista y recordar que cada candidato es único. Debe evitar las preguntas cerradas con el fin de evitar acaparar la conversación y que el alumno se limite a responder tan sólo *sí* o *no* a sus cuestionamientos. Por último, pero no menos importante, para llegar a resultados válidos y confiables debe saber qué preguntar para conseguir que la entrevista cubra los contenidos lingüísticos esperados.

Tomemos en cuenta que la problemática de las entrevistas recae en el tipo de puntuación requerida, la cual, si no se tiene el cuidado, puede llegar a ser absolutamente subjetiva. Es evidente que los juicios de valor que los entrevistadores deben emitir son mucho más difíciles que los de las pruebas objetivas en las que únicamente se requiere una clave que nos indique si la respuesta es correcta o incorrecta.

Como se comentó previamente al mencionar el tipo de pruebas subjetivas, para que un examinador tenga un mejor desempeño en este tipo de tareas, debe poder contar con una *escala de valoración* que le permita referirse, a través de *descriptores* establecidos, al nivel de competencia y actuación lingüística del candidato.

Alderson, et al (1998) comentan sobre dos tipos de escalas de valoración (aunque cada institución o experto en evaluación puede emitir la propia):

- La escala *holística* (también llamada *global* o de *impresión general*) con la que el evaluador emite un juicio de eficacia general sin tomar en cuenta detalles concretos en la producción. En este tipo de escalas los descriptores se limitan a términos como “Excelente” (Inglés natural con errores mínimos y realización completa de la tarea encomendada), “Muy bien” (Más que un conjunto de frases simples, con buen vocabulario y estructuras. Algunos errores básicos), “Bien” (Realización simple aunque precisa de la tarea encomendada, con algunos errores serios. “Suficiente” (Razonablemente correcta aunque torpe y no comunicativa; tratamiento natural y justo del tema con algunos errores), “Pobre” (Vocabulario y gramática no adecuados al tema). “Muy pobre” (Incoherente. Los errores muestran la ausencia de conocimientos básicos de inglés). (De: *UCLES International Examinations in English as a Foreign Language General Handbook*, 1987, en Alderson, et al (1998).
- La escala analítica, en la que se deben tomar en cuenta otros componentes de la actuación lingüística como la caligrafía, organización de párrafos, gramática, vocabulario, etc. Las puntuaciones de los candidatos pueden ser distintas para cada componente y promediarse posteriormente para obtener un resultado final. En este tipo de escalas los descriptores son mucho más precisos. Por ejemplo:

Tabla 2. Escala de valoración Analítica

<u>Relevancia y adecuación del contenido</u>	
0.	La respuesta no tiene prácticamente ninguna relación con la tarea encomendada. Respuesta totalmente inadecuada.
1.	Respuesta de relevancia limitada respecto a la tarea. Existen posiblemente lagunas importantes en el tratamiento del tema y/o repetición inútil.
2.	Responde en gran parte a la tarea, aunque puede haber algunas lagunas o información redundante.
3.	Respuesta relevante y adecuada a la tarea.
	[...]
<u>Gramática</u>	
0.	Casi todas las estructuras gramaticales son incorrectas.
1.	Frecuentes incorrecciones gramaticales.
2.	Algunas incorrecciones gramaticales.
3.	Casi no se encuentran incorrecciones gramaticales.

Tomado de: *Test of English for Educational Purposes*, Associated Examining Board, UK, 1984, en Alderson, et al (1998).¹⁴

¹⁴ Nota: El cuadro presenta tan sólo dos ejemplos del tipo de componentes evaluados en las escalas analíticas. Para mayores detalles referirse a Alderson, et al, 1998.

Como se puede observar, es fundamental revisar y cumplir con los criterios que requiere la evaluación oral para poder emitir resultados que sean no solo éticos y justos, sino válidos y confiables para quienes ponen en nuestras manos decisiones trascendentales que pueden afectarles o beneficiarles de manera personal.

III. MARCO CONTEXTUAL

▪ Descripción del Centro de Idiomas, del examen escrito y la entrevista oral.

El caso que aquí nos ocupa es el del proceso de clasificación que se aplica a los interesados en ingresar a los curso de inglés del Centro de Idiomas de la Universidad Iberoamericana. Es un proceso que incluye un examen escrito y una entrevista en LM y que tiene por objetivo determinar el nivel en el que un estudiante debe iniciar sus estudios dentro del programa. Tomar una decisión acertada es de suma importancia para el adecuado desarrollo del Centro, de los programas, y principalmente, del estudiante.

El Centro de Idiomas de la Universidad Iberoamericana, Cd. de México, fundado en los años 60, ha ofrecido desde entonces la enseñanza de idiomas como el inglés, francés, alemán, italiano, y recientemente portugués y chino mandarín. Las clases de lenguas están dirigidas tanto a alumnos de la universidad como al público en general.

En el caso particular del idioma inglés, la universidad ha establecido que es requisito que los estudiantes que terminan la licenciatura hayan cumplido, de preferencia a más tardar durante el tercer semestre de su licenciatura, con un examen que compruebe su nivel de dominio del idioma. Uno de los exámenes aceptado en la actualidad por la UIA es el *Test of English as a Foreign Language* (TOEFL), ya sea en su modalidad de examen sobre papel (*paper based*) aplicado a nivel Institucional dentro de la misma universidad, el cual debe ser aprobado con un mínimo de 500 puntos (de un total de 677), lo que equivale a un nivel intermedio; o en su modalidad por internet (el iBT o *Internet Based Toefl*), con un puntaje equivalente a 61 puntos para alumnos que hayan iniciado sus estudios antes de 2008. Quienes lo hayan hecho posteriormente, deben realizar el examen iBT con una puntuación de 80, equivalente a 550

puntos del examen Institucional. Asimismo la universidad acepta las equivalencias de otros exámenes reconocidos internacionalmente como el IELTS y los de la universidad de Cambridge.

A lo largo de los años el Centro de Idiomas ha ido evolucionando tanto en técnicas de enseñanza como en programas de estudio. Hoy en día se ofrecen para el idioma inglés seis niveles integrados por dos cursos cada uno de ochenta horas, lo que equivale a 960 hrs. de estudio para cubrir el programa completo¹⁵.

Tabla 3. Panorama de cursos impartidos por la UIA (2007).

Principiante	Elemental	Pre-Intermedio	Intermedio	Intermedio Alto	Avanzado	Total
2 cursos: 160 hrs.	2 cursos: 160 hrs	2 cursos: 160 hrs	2 cursos: 160 hrs	2 cursos 160 hrs	2cursos: 160 hrs	6 niveles: 12 cursos 960 hrs.

Adicionalmente, el Centro ofrece el curso de preparación para el examen TOEFL (al que un alumno puede integrarse si cuenta con aproximadamente 640 hrs. de estudio, esto es, al terminar el nivel intermedio). Asimismo, se ofrecen cursos de conversación para diferentes niveles, así como otros especializados como Inglés para Negocios, Inglés para Abogados y cursos especializados como *Reading* o *Writing for professional purposes* (la duración de éstos puede variar entre 60 y 80 hrs.)

En el Centro de Idiomas también se ofrecen opciones más especializadas a través de dos diplomados: Traducción Inglés-Español, y Formación de profesores en la enseñanza del español como lengua extranjera.

Es evidente que al ser una alternativa de estudio abierta a todo público y con una amplia gama de ofertas educativas en la enseñanza-aprendizaje del idioma, el tipo de aspirantes a

¹⁵ La distribución de cursos y habilidades a cubrir en la UIA con respecto al idioma inglés está basada en los niveles establecidos por el Marco de Referencia Europea, cuyos niveles de competencia también son seis, los cuales equivalen a los establecidos en la UIA como sigue: A1: *break through level*, A2: *Level 1*, B1: *Level 2*, B2: *Level 3*, C1: *Level 4*, C2: *Level 5*. Como se ha establecido previamente, uno de los objetivos principales del Marco es establecer niveles comunes de competencia en distintos idiomas con el fin de promover la certificación en Europa de los interesados con respecto a sus habilidades en alguna lengua en particular.

ingresar a los cursos puede provenir de diversos contextos escolares y situaciones particulares, esto es: de escuelas tanto públicas (con pocas horas de inglés a la semana a partir de la educación secundaria) como de escuelas privadas (donde se enseña el idioma desde preescolar, y donde incluso muchas de las materias se imparten en la segunda lengua); así mismo, los interesados pueden ser personas que cuentan con un capital cultural que les ha permitido estar en contacto con el inglés a lo largo de su vida, y otras que no han tenido dicha oportunidad.

Por lo anterior, ante la notable variedad de conocimientos de quienes solicitan ingreso, establecer el nivel de dominio del idioma de cada aspirante es fundamental para poder ofrecer la mejor alternativa académica dentro de los programas de inglés ofrecidos por la UIA y para el buen funcionamiento del Centro de Idiomas.

Hemos comentado que para detectar los conocimientos de un estudiante, se requiere un proceso de clasificación que contemple un examen que ayude a determinar su manejo de la lengua en cuestión. En el caso de la UIA, durante muchos años el Centro de Idiomas intentó emplear este tipo de examen, para lo cual se utilizaban copias (a veces de poca calidad) de ejercicios de textos, o el examen final de un programa, u otro tipo de pruebas que no habían sido creados para evaluar el nivel de conocimiento de los alumnos. Además, la presentación era inadecuada y los resultados poco satisfactorios.

En 1999 esta situación cambió puesto que se creó un instrumento que tuviera el objetivo descrito. El examen fue elaborado por una de las profesoras que trabajaba en el Centro en ese entonces y revisado por nativo parlantes de la lengua. Este examen se diseñó con base en la estructura de los programas que se ofrecían en ese entonces, con una duración de 90 horas cada uno, haciendo un total de 1,080 hrs. de estudio. Cabe mencionar que el examen nunca fue revisado científicamente a través de la estadística o de las pruebas pertinentes para el buen

funcionamiento de los reactivos, y aunque se presentaban problemas en la ubicación de los estudiantes, la falta de tiempo y de personal académico no permitía ahondar en el asunto.

En el año 2003 el Centro decidió actualizar el programa y los libros de texto y la autora de este trabajo comenzó a observar que se presentaban dificultades al respecto. Por lo tanto, con el fin de mejorar la validez y confiabilidad de la prueba, se empezaron a realizar modificaciones a la versión original. Inicialmente se revisaron los reactivos para que tuvieran un orden progresivo con respecto al conocimiento presentado en cada curso. Asimismo se trabajó para detectar el funcionamiento de las alternativas a elegir (al ser un examen de opción múltiple) a través de ciertas pruebas con SPSS y se le pidió a los profesores revisar los reactivos y alternativas hasta que se llegó a una nueva versión a principios de 2006. Hasta finales de ese año no se había podido realizar un análisis detallado de reactivos (coeficiente de discriminación e índice de dificultad) por no contar con el número de pruebas suficiente para obtener resultados adecuados.

El examen escrito que se utiliza desde primavera de 2006, tiene las siguientes características y cuenta con los siguientes elementos:

- ❑ Consta de un cuadernillo que cubre desde el nivel Principiante hasta el Avanzado
- ❑ En cada página hay 15 preguntas correspondientes a cada curso (un total de 180 reactivos).
- ❑ Es de opción múltiple. Cada pregunta tiene 4 alternativas a elegir.
- ❑ La hoja de respuestas es independiente al cuadernillo.
- ❑ Se califica manualmente a través de una plantilla.
- ❑ Cuenta con una tabla para determinar el nivel correspondiente.

Adicionalmente, se realiza una evaluación oral para precisar el resultado que el examen escrito arroja. La entrevista tiene las siguientes características:

- ❑ Se realiza de manera individual.
- ❑ Tiene una duración aproximada de 5 a 10 minutos por candidato, mediante la cual se detecta conocimiento e intereses particulares del aspirante entrevistado.
- ❑ Se cuenta con una guía de conocimiento por nivel que permite a los profesores saber qué temas lingüísticos el alumno debe manejar en cada curso, con el fin de asegurar que los maneje de manera adecuada.

Una vez evaluadas la parte escrita y la oral, los profesores que realizan los exámenes de colocación establecen el nivel definitivo en el cual ubicar a cada estudiante. En caso de que haya discrepancias entre los resultados de una y otra sección, se ubica al interesado en un nivel intermedio entre ambos con el fin de que mediante el curso correspondiente, el estudiante consiga equilibrar sus conocimientos del idioma en ambas áreas.

Es importante mencionar que uno de los objetivos de los programas de inglés ofrecidos por el Centro de Idiomas es lograr que los estudiantes que ingresaron antes de 2008 a la UIA avancen de manera gradual para lograr llevarlos al nivel requerido para el curso de preparación para el examen TOEFL Institucional (o que quienes ya tienen el conocimiento necesario se inscriban directamente a este programa). Esto con el fin de proporcionarles una opción adicional para aprobar el requisito del idioma a través de este curso. Por lo tanto, en el examen de clasificación se le da mayor peso al manejo de estructuras que a la expresión oral (aproximadamente un 60% - 40%, aunque esto no ha sido establecido formalmente). Esto se debe a que la sección gramatical del examen TOEFL, y por lo tanto del curso, es muy demandante, además de que en el examen Institucional de hecho no existe una sección que evalúe la habilidad oral.

Un estudiante puede comunicarse de forma oral de manera efectiva, pues aún cuando cometa errores, su lenguaje corporal y el contexto en el que se encuentra pueden ayudarle a

transmitir sus ideas. Sin embargo, si no tiene la base estructural, le será difícil comprender lecturas, interacciones auditivas y sobre todo, realizar ejercicios gramaticales específicos para una prueba escrita.

Antes de utilizar la última versión del examen, se presentaban mayores discrepancias en la colocación de los interesados en los cursos. Hoy en día, esto ha disminuido bastante, aunque aún se presentan casos de una mala ubicación, los cuales pueden ocurrir en las siguientes circunstancias:

- ❑ En ocasiones, una vez iniciado el curso, los profesores se percatan de que algún estudiante (algunas veces más de uno) tiene un nivel de dominio del inglés (más alto o más bajo) que no corresponde al del nivel del curso en el que fue ubicado, por lo que se procede a cambiar al estudiante al curso conveniente. A veces esto ocurre cuando ya ha pasado una semana del inicio de clases.
- ❑ En otras ocasiones, el alumno mismo informa a la coordinación o al profesor que no se siente bien ubicado en el nivel al que asiste, porque tiene menor o mayor conocimiento del idioma, por lo que también se procede al cambio de nivel, una vez consultado el profesor asignado.

En cualquiera de los casos anteriores, si la clase no estuviera disponible, se reubica al alumno en la clase inmediata inferior y se le explica la conveniencia de hacerlo así. Algunos de ellos prefieren quedarse en la clase de origen, comprometiéndose a hacer un mayor esfuerzo para completar el curso con éxito.

En las situaciones más desafortunadas, el maestro no alcanza a darse cuenta a tiempo del problema ni el alumno lo externa, por lo que éste simplemente, deserta.

En resumen, aún cuando la sección escrita del examen ha sido mejorada y revisada con mayores elementos y no obstante que la sección subjetiva del examen, esto es, la entrevista, cuenta ahora con una guía para los profesores con respecto al nivel de dominio que deberían sustentar los estudiantes en cada curso, se desea conocer en qué medida el proceso de clasificación como se maneja actualmente es acertado para ubicar a los aspirantes en el curso que realmente les corresponde con base en su habilidad lingüística.

IV. METODOLOGÍA

▪ **Tipo de investigación.**

El enfoque del presente trabajo puede considerarse como *mixto-dominante* (Hernández, Fernández, Baptista, 2003), ya que, aunque prevalece la perspectiva cualitativa basada en la observación y descripción del proceso así como en la recolección de datos tomados de la opinión de los participantes sobre el proceso y su resultado, también se utilizan componentes de la investigación cuantitativa como la estadística, con el fin de analizar el grado de confiabilidad y de evidencia de validez del instrumento escrito utilizado en el proceso de clasificación, así como para analizar las respuestas que arrojaron los distintos cuestionarios que se aplicaron a los participantes.

▪ **Diseño.**

La estrategia utilizada para alcanzar los objetivos del estudio fue a través de un diseño *no experimental* (Hernández, Fernández, Baptista, 2003) en el que se observó el fenómeno que nos ocupa tal y como se presenta en su contexto natural.

A su vez, el diseño tuvo una dimensión transversal-descriptiva. Esto es, las observaciones y recolección de datos se llevaron a cabo para un momento único en el tiempo. Se recolectaron datos descriptivos basados en la observación del proceso de clasificación del idioma inglés para ingresar a los cursos en el periodo específico de Primavera de 2007.

Por otro lado, el diseño se basó en la información estadística que arrojó el examen escrito con base en la recolección de datos de las respuestas de los candidatos que realizaron el proceso en los periodos de ingreso de Primavera, Verano y Otoño de 2006, y Primavera de 2007. La obtención de estos datos tomó mayor tiempo con el fin de obtener información suficiente para poder realizar el análisis de los reactivos de la prueba escrita.

▪ Unidad de análisis (muestra)

Con base en Hernández, Fernández y Baptista (2003), la unidad de análisis de esta investigación (es decir, el “objeto” o “muestra” a estudiar) puede considerarse un estudio de caso¹⁶. Los datos de la muestra para el análisis del proceso de clasificación procedieron de sujetos voluntarios interesados en conocer su nivel del idioma inglés, para en determinado caso, inscribirse a los cursos impartidos en la UIA en Primavera de 2007. Estos sujetos fueron alumnos, ex-alumnos, personal de la universidad, o candidatos externos de 17 años en adelante, de cualquier sexo, carrera de estudio, ocupación, o profesión. La muestra incluyó también a personal de la empresa Bimbo interesado en participar en los cursos ofrecidos en sus propias instalaciones.

Una vez iniciados los cursos, a la muestra se sumó la información que arrojó el cuestionario entregado a los profesores sobre el desempeño de los alumnos, para conocer si desde su punto de vista había alumnos mal clasificados.

▪ Instrumentos de recolección de datos

Para la descripción del periodo de clasificación de P-07, se utilizaron los siguientes instrumentos:

- Hoja de información del aspirante.¹⁷ Tuvo por objeto recabar información general de quien realizaría el examen, tal como datos personales (nombre, teléfono, edad); si era alumno o trabajador de la UIA, o candidato externo; las razones para realizar el proceso de clasificación (requisito de idioma de la UIA; superación personal o laboral); trayectoria escolar (de primaria a profesional, y si realizó los estudios en escuela pública o privada); y fecha aproximada de sus últimos estudios de inglés. Todos estos datos con

¹⁶ No manejan el término Estudio de Caso como algunos otros autores (Grinell 1997) que lo consideran como metodología o diseño.

¹⁷ Apéndice 1.

el fin de conocer los antecedentes de los aspirantes a los cursos, y comenzar a recolectar información para posibles futuras investigaciones. Un elemento más que incluyó el instrumento y que sirvió como referencia para comparar el curso en el que se clasificó al candidato, fue solicitarle determinar cual consideraba que era su nivel de idioma en el momento de realizar el proceso, para lo que se le pidió basarse en un cuadro adaptado por la autora de este proyecto tomado del MCREL (2002), considerando la teoría de Le Blank y Pinchaud (en Murray 1999) relacionada con la relevancia de la autoevaluación.

- Cuestionario de salida para aspirantes.¹⁸ Tuvo como objetivo conocer sus impresiones con respecto al proceso de clasificación que realizó, para poder determinar la validez aparente (“la credibilidad aparente o aceptación pública de la prueba”, Ingram, 1977) tanto del proceso en general, como del instrumento escrito. Se proporcionó una vez que se le informó el curso correspondiente.
- Cuestionario para profesores.¹⁹ Se aplicó dos semanas después de iniciados los cursos con el fin de conocer si algún profesor detectó a algún alumno cuyo nivel no fuera el del curso correspondiente. En caso de ser así, se procuró dar atención al caso, reubicar al alumno si era necesario, y darle seguimiento a su desempeño.
- Cuestionario para alumnos.²⁰ Se entregó a los alumnos inscritos al mismo tiempo que a los profesores para confirmar si el proceso de clasificación los había ubicado en el curso apropiado o si se requería algún cambio de nivel.
- Grabación de veintitrés evaluaciones orales para conocer cómo se llevan a cabo las entrevistas.²¹

¹⁸ Apéndice 2.

¹⁹ Apéndice 3.

²⁰ Apéndice 4.

²¹ Resultados resumidos más adelante. No se incluye en anexos.

- Examen escrito. Se capturaron las respuestas de cada alumno que realizó el examen, compuesto de 180 reactivos y 4 alternativas de opción múltiple para poder realizar un análisis de reactivos que indicara qué ítems funcionaban adecuadamente, cuales deberían corregirse, y/o cuáles eliminarse para hacer más confiable la prueba.²²

▪ Fases de la investigación

- Inicios

Es importante recordar que el problema surge en 2003 con el cambio de los programas para los cuales se había diseñado el primer examen específicamente escrito para el Centro de Idiomas que se utilizó antes de ese año y que se realizó para otros programas. A principios de 2004, al detectar inconsistencias en la clasificación de los estudiantes, la autora de este trabajo comienza a realizar correcciones y cambios al instrumento.

Durante esta etapa:

- Se revisaron los reactivos y las 4 opciones a elegir en varias ocasiones a nivel de lenguaje.
- A finales de 2004 se pidió apoyo a los profesores del idioma que se encontraban impartiendo algún curso. Se les solicitó:
 - responder el examen.
 - hacer comentarios sobre los aspecto del lenguaje que consideraban evaluaba cada reactivo e indicar si los temas coincidían con el programa impartido.

²² No se incluye en anexos para asegurar la confidencialidad del examen.

- hacer sugerencias en caso de que consideraran la necesidad de hacer algún cambio, ya fuera al reactivo o a las opciones, con la intención de validar el contenido a través de expertos.
- evaluar la validez aparente del instrumento, para indicar si en esencia consideraban que el formato del examen era adecuado o no, de acuerdo con el propósito con el que fue creado.
- Durante 2005 se realizaron algunos cambios y se elaboró la Matriz-Contenido-Conducta y la Tabla de Especificaciones correspondientes²³ que, con base en los programas utilizados, permitieron realizar cambios al orden de los reactivos, agregar temas faltantes y eliminar temas repetidos. La Matriz permitió también elaborar las especificaciones para la evaluación oral con el fin de que los profesores cuenten con una base que les indique el conocimiento que se espera maneje el candidato.
- Asimismo, se re-diseñó e incluyó la sección Avanzada del examen, y se solicitó la revisión de los reactivos a algunos de los profesores nativo-hablantes de inglés. De tal modo, se obtiene la versión utilizada actualmente.

- Desarrollo

En Primavera de 2006 se comenzó a aplicar a los aspirantes a los cursos la última versión del instrumento escrito y a capturar la información en SPSS con el fin de realizar un análisis de reactivos e intentar comprobar su confiabilidad y determinar su evidencia de validez. Se continuó recabando información durante todo el año en virtud de que se requería obtener suficientes datos.

La revisión de la literatura ayudó a determinar que el instrumento escrito no es el único elemento que puede influir en una correcta o inadecuada clasificación del nivel del idioma, por lo

²³ Apéndices 5 y 6.

que se tomó la decisión de evaluar el proceso de clasificación completo, y no sólo la prueba de opción múltiple.

Así surgió la pregunta de investigación: ¿Cuál es la efectividad en cuanto a su confiabilidad y evidencia de validez del proceso de colocación que se lleva a cabo en el Centro de Idiomas de la UIA?

A finales del ciclo escolar de Otoño de 2006 se piloteó un cuestionario para profesores con el fin de conocer si en ese periodo habían detectado a algún estudiante cuyo nivel no fuera el adecuado para el curso correspondiente. Asimismo se elaboró otro cuestionario para estudiantes para conocer su opinión con respecto a su nivel, y sobre el proceso de clasificación del idioma, sin importar cuándo lo hubieran realizado (ya que algunos de los estudiantes tenían ya varios ciclos de estudio en el Centro de Idiomas, y no necesariamente habían realizado el proceso de clasificación recientemente.)

A continuación se resumen los resultados obtenidos de los cuestionarios aplicados en Otoño 2006:

Tabla 4. Cuestionario para profesores (Piloto, Otoño 2006)

No. de cursos	Total de alumnos inscritos	Total de alumnos que respondieron el cuestionario	No. de alumnos <u>mal clasificados</u> de acuerdo con los profesores	No. de alumnos con nivel más bajo	No. de alumnos con nivel más alto
7	80	70	11 = 15.7%	5	6

Tabla 5. Cuestionario para alumnos (Piloto, Otoño 2006)

Alumnos que respondieron el cuestionario:	70	Comentarios
Alumnos que se consideraron bien clasificados cuando realizaron el proceso	67 = 95.7%	
Alumnos que consideraron que no estuvieron bien clasificados cuando realizaron el proceso	3 = 4.3%	
Alumnos que consideraron que no hubieran obtenido el mismo resultado si hubieran realizado el proceso al siguiente día	13 = 18%	3 no indican la causa. 1 por falta de tiempo. 4 al conocerlo estarían menos nerviosos. 1 porque tenía sueño. 1 por falta de concentración.

		1 porque hubiera estudiado. 2 porque se sintieron presionados.
Alumnos que consideraron que al participar en el curso su nivel no fue el adecuado	7 = 10%	6 consideraron que su nivel era más alto en distintas áreas como oral, y lectura. 1 en sus primeros 2 cursos se sintió bien ubicado. En los últimos 3 sintió que había gente con mayor nivel. Manifestó problemas de comprensión auditiva.
Alumnos que fueron cambiados de nivel	1	Expresó su inquietud. Se evaluó su examen escrito y se le invitó a participar durante un par de clases en el nivel inmediato superior. Tanto la alumna como la profesora confirmaron que podía participar ahí. La alumna reportó su satisfacción con el cambio
Alumnos que NO fueron cambiados de nivel	6	3 No informaron. Prefirieron repasar. 2 No informaron. 1 Expresó su inquietud. Se evaluó su examen escrito, el cual sólo había completado hasta el nivel en el que se le ubicó. Se le realizó otro examen y se determinó que sus estructuras eran muy deficientes para alguien cuyo objetivo era llegar a la preparación del examen Toefl. Se le sugirió repasar, pero no quedó satisfecho.

Después de analizar las preguntas de los cuestionarios y la información que arrojaron y realizar cambios mínimos, se decidió mantenerlos para su futura aplicación.²⁴

Asimismo, se hicieron los preparativos para llevar a cabo la observación específica del proceso de clasificación a ocurrir en Primavera 2007 y aplicar los diversos cuestionarios para conocer la opinión de los usuarios (examinadores, candidatos, profesores y estudiantes) sobre el proceso de clasificación en su conjunto.

- El proceso de clasificación de Primavera 2007

El proceso de clasificación de los aspirantes a ingresar a los cursos de inglés ofrecidos por la UIA se llevó a cabo del 5 al 17 de enero en los horarios previamente establecidos de lunes, miércoles y viernes a las 10 ó a las 16 hrs.

²⁴ Estos son el Cuestionario para Profesores, y el Cuestionario para Alumnos, anexos 3 y 4 a los que se hace referencia en la página 62.

Los interesados no requieren hacer cita. Simplemente se presentan en el edificio del Centro de Idiomas con su recibo de pago 5 minutos antes de la hora, y se les indica en qué salón se lleva a cabo el proceso.

El primer día se solicitó únicamente a una profesora para atender a los candidatos. Al turno de la mañana se presentaron 4 personas. La profesora dio las instrucciones en inglés de cómo completar el examen escrito, haciendo énfasis en que lo hicieran hasta donde les fuera posible y se sintieran seguros. Asimismo explicó la forma de llenar la hoja de datos personales. Sin embargo algunos alumnos no parecieron entender las instrucciones, por lo que la profesora tuvo que hacer algunas aclaraciones de manera individual posteriormente.

Al concluir el proceso se le recordó a la maestra repasar las instrucciones en español para asegurar que la información quedara clara. El hecho de que la maestra diera las instrucciones en inglés únicamente pudo haber sido causa de nerviosismo de su parte al haber un observador en el aula, pues indicó que normalmente sí lo hace de esa manera. La profesora es una persona con varios años de experiencia en el proceso de clasificación de la UIA.

Al examen de las 16 hrs. se presentaron 6 alumnos. La profesora dio instrucciones tanto en inglés como en español y mencionó que sólo tenían que completar hasta donde pudieran.

Una vez que el alumno concluye o decide detenerse en la evaluación escrita, hace entrega a la profesora, quien lo califica con la plantilla. La maestra determina el nivel escrito con la escala correspondiente, y posteriormente lo entrevista para confirmar a través de la producción oral el nivel en el que podrá inscribirse. Una vez terminada la entrevista, y habiendo informado al candidato su nivel, se le solicitó completar el cuestionario de salida, en el que se les pide comentar su opinión sobre el proceso y si está de acuerdo con el nivel que se determinó.

El segundo día se solicitó a 2 maestras para cada turno, al cual se presentaron 8 alumnos por la mañana y tan sólo 2 en la tarde. Se repasó con las profesoras el proceso antes del examen,

y esta vez no se presentó ninguna duda por parte de los interesados ya que, además de dar las instrucciones en inglés, se aseguró que quedaran claras en español.

El tercer día, enero 10, se presentaron 8 alumnos por la mañana y 16 por la tarde, mientras que el viernes 12, el cuarto día, hubo 24 candidatos en la mañana y 18 por la tarde. Para las últimas 2 fechas oficiales se citó a 3 profesoras pues es evidente que la demanda va creciendo conforme pasan los días y se acerca el cierre de inscripciones. Por esta razón se amplió el periodo de exámenes 1 día más.

En resumen, durante el procedimiento de clasificación de nivel se llevó a cabo lo siguiente:

- Se dieron instrucciones en español y en inglés a los candidatos para indicarles cómo llenar la hoja de datos personales, y la manera de auto-evaluar su nivel de conocimiento de acuerdo a la tabla basada en el MERCL incluida al reverso de la mencionada hoja de datos personales.
- Se les explicó que una vez completada la hoja de datos personales podían proceder a realizar el examen, contestando columna por columna y teniendo el cuidado de asegurarse de que estuvieran trabajando en la sección correcta.
- Se les indicó que únicamente hay una posible respuesta para cada pregunta.
- Se les recomendó realizar hasta donde se sintieran confiados, y se hizo hincapié en que no era necesario completar el examen.
- Se les informó que una vez que terminaran la sección escrita, pasarían a una entrevista y que al finalizar se les informaría el curso a inscribir.

De este modo, una vez que el candidato entrega el examen escrito, debe esperar unos minutos a que éste sea revisado para posteriormente realizar la entrevista. La ventaja de contar con dos o tres profesoras permite que, en días muy saturados, una de ellas revise la prueba escrita para establecer el nivel de habilidad respectivo y hacer las recomendaciones pertinentes a la o las

entrevistadoras para que éstas puedan tomar la mejor decisión sobre el nivel definitivo a través de la entrevista oral.

Las entrevistas se realizaron de manera individual en salones contiguos al del examen escrito. Como se comentó previamente en el marco contextual, la problemática de este tipo de evaluación subjetiva es que cada entrevistador puede tener un criterio distinto para la toma de decisiones. Por lo tanto, con el objeto de contar con el menor número de discrepancias, se procuró que fueran dos las profesoras que realizaran las entrevistas. Sin embargo las necesidades de horario de cada una y el número tan elevado de candidatos a evaluar en el periodo, requirió que una profesora más interviniera. Asimismo, con el fin de que el criterio de decisión fuera lo más confiable posible, se le pidió a la profesora a cargo de la entrevista que fuera quien determinara el curso definitivo e indicara el criterio utilizado para ello.

En los tres casos son maestras con mucho tiempo de experiencia no sólo impartiendo clase, sino aplicando y calificando el examen escrito así como realizando las evaluaciones orales del proceso de clasificación, actividad que han venido realizando cada periodo de inscripciones desde hace varios años. Todas ellas conocen bien la mayoría de los programas y los han impartido en distintas ocasiones.

Con el fin de conocer de manera más precisa la forma en que llevan a cabo esta parte del proceso, y con la previa autorización de los candidatos antes de hacerlo, se grabaron al azar 23 entrevistas.

Algunos elementos a destacar de lo que ocurre en cada una de ellas es la amabilidad y disposición de las entrevistadoras, así como la confianza que procuran transmitir al candidato y la buena relación que se establece entre maestra-alumno. Por lo general hay risas y comentarios personales que permiten observar estos factores en cada una de ellas.

Todas las entrevistas inician con un saludo y preguntando el nombre del candidato. Las profesoras inician la interacción basadas en los primeros datos que reciben en cuanto al examen escrito, pero adaptan el nivel del idioma de acuerdo con el grado de producción que detectan en las primeras expresiones del estudiante. La duración varió entre 3 y 8 minutos, aunque hubo una que duró 14.

En la mayoría de las entrevistas se detecta por parte de las profesoras cierta similitud en los temas a tratar, como son:

- descripción de la familia
- si es estudiante, la carrera que estudia
- la razón de elegirla
- la descripción de algún evento del pasado que el alumno mencione
- dónde ha estudiado inglés
- el curso que le interesa
- si trabaja, dónde y qué actividades realiza
- sus planes para el futuro

La profesora 1 grabó 7 entrevistas. Utilizó sólo el idioma inglés y se limitó a estructuras muy generales como presente, pasado, futuro y presente perfecto.

La profesora 2 grabó 14 entrevistas. Su estrategia parece muy estructurada y enfocada al nivel del estudiante, ya que una vez que detecta el grado de producción del alumno, sus intereses, antecedentes familiares, etc. les pide traducir del español al inglés ciertas estructuras como el uso de condicionales, auxiliares, expresiones, etc. En la entrevista se escucha cómo le pide tiempo al

candidato para revisar el tipo de pregunta que debe hacerle para verificar su uso de la misma y poder determinar el nivel y curso correspondiente.

La profesora 3 grabó sólo 2 entrevistas. En general revisó los mismos temas y utilizó los mismos tiempos básicos que las otras dos (presente, pasado, futuro, presente perfecto). Utilizó principalmente inglés para especificar situaciones que le permitieran al alumno producir estructuras específicas como condicionales, sugerencias, prohibiciones, etc. Sólo en algún caso extremo o para asegurarse de obtener la estructura que pretendía que el alumno produjera, recurrió al español.

Al terminar el procedimiento:

- Se le informó al candidato su nivel y el curso que le correspondía.
- Se le pidió completar el cuestionario de salida para conocer su percepción del proceso.
- Se le dio información relacionada con el procedimiento para inscribirse al curso.

V. ANÁLISIS DE RESULTADOS

▪ **Reporte cualitativo de resultados del proceso de clasificación.**

Es importante resaltar que, el número de evaluaciones de Primavera 2007 fue extraordinariamente alto (249 estudiantes), comparado con los distintos periodos en 2006, en los que se evaluó a un total de 309 estudiantes (116 en primavera, 39 en verano, 154 en otoño). La razón fue que, como ya se mencionó en la descripción de la muestra, la empresa Bimbo solicitó evaluaciones extras para abrir cursos en sus nuevas instalaciones de Tecnoparque, lo que fue muy positivo para el objetivo de este proyecto. Estas evaluaciones se llevaron a cabo los días 13, 20 y 27 de febrero. En este proceso de clasificación también participaron entre 2 y 3 profesoras en cada ocasión bajo el mismo esquema.

Como se observa en la siguiente tabla, de los 249 estudiantes evaluados el 65.1 % se inscribió a los cursos, ya fuera en la UIA, o en sedes externas de la empresa Bimbo

Tabla 6. Candidatos que se inscribieron

		Frequenc	Percen	Valid	Cumulativ Percen
Valid	sí	162	65.1	65.1	65.1
	no	87	34.9	34.9	100.0
	Total	249	100.0	100.0	

La distribución de los estudiantes inscritos en las distintas sedes quedó de la siguiente manera: 76 en la UIA, 17 en las oficinas corporativas de la empresa mencionada, y 69 en sus nuevas instalaciones.

Tabla 7. Lugar de inscripción

	Frequenc	Percen	Valid	Cumulativ Percen
Valid Uia	76	30.5	30.5	30.5
Corp	17	6.8	6.8	37.3
TP	69	27.7	27.7	65.1
Ningun	87	34.9	34.9	100.0
Total	249	100.0	100.0	

Los alumnos fueron clasificados en los distintos niveles y cursos impartidos por la universidad, como puede observarse en la tabla siguiente.

Tabla 8. Curso a inscribir

	Frequenc	Percen	Valid	Cumulativ Percen
Valid Ingles 1	40	16.1	16.1	16.1
Ingles 2	13	5.2	5.2	21.3
Ingles 3	30	12.0	12.0	33.3
Ingles 4	18	7.2	7.2	40.6
Ingles 5	27	10.8	10.8	51.4
Ingles 6	24	9.6	9.6	61.0
Ingles 7	26	10.4	10.4	71.5
Ingles 8	16	6.4	6.4	77.9
Ingles 9	20	8.0	8.0	85.9
ingles 10	12	4.8	4.8	90.8
Ingles 11	5	2.0	2.0	92.8
ingles 12	1	.4	.4	93.2
Toefl	17	6.8	6.8	100.0
Total	249	100.0	100.0	

A continuación se muestra el número de alumnos que se inscribieron en cada curso:

Tabla 9. Curso a inscribir / No. inscritos: Crosstabulation

Count		¿Se inscribió?		Total
		sí	no	
Curso a inscribir	Ingles 1	29	11	40
	ingles 2	7	6	13
	Ingles 3	20	10	30
	Ingles 4	14	4	18
	Ingles 5	20	7	27
	Ingles 6	11	13	24
	Ingles 7	20	6	26
	Ingles 8	7	9	16
	Ingles 9	16	4	20
	ingles 10	3	9	12
	Ingles 11		5	5
	ingles 12		1	1
	Toefl	15	2	17
Total	162	87	249	

Cabe comentar que en las instalaciones de la universidad se cancelaron los cursos 2, 10, 11, 12 por no cumplir con el número de alumnos requerido para abrirlos, el cual es de un mínimo de ocho.

En el caso del curso 2, sólo se inscribieron 2 alumnos en la UIA, los otros 5 lo hicieron en sedes externas. Los 2 alumnos del nivel 2 de la UIA decidieron participar en el curso 1. Un alumno clasificado en el curso 2 en sede externa participó en el curso 1. Así lo hizo 1 clasificado en curso 3. En estos dos casos no pudo determinarse si se equivocaron de curso o si decidieron hacerlo con plena conciencia. Sin embargo cuando la situación se detectó, se habló con el profesor para ver si no le causaba problema a él, al grupo correspondiente, o a estos dos alumnos. El maestro habló con ellos y le indicaron sentirse bien trabajando en ese grupo. El profesor confirmó que podían trabajar en el curso al que asistían.

Una persona ubicada en el curso 4 participó en el curso 1 debido a que no se abrió el que le correspondía en la sede externa en la que labora. Asimismo, alguien clasificado en el curso 3 tuvo que hacerlo en el 4 por falta de grupo en la sede de su lugar de trabajo. En este caso a esta

persona se le permitió subir de nivel en virtud de que anteriormente ya había trabajado en este curso. Dejó el idioma por un tiempo, y aunque su nivel bajó, se consideró que podría repetirlo.

Del mismo modo, hubo una persona clasificada en el curso 4 en la UIA, quien después de 1 semana de curso habló con la profesora y con la coordinación del Centro de Idiomas para expresar que sentía que su nivel era más alto. Se le invitó a participar en el curso 5 y tanto la maestra de ese grupo como ella misma confirmaron que era mejor que trabajara en éste.

En el caso del nivel 10 en la UIA, 1 alumno se incorporó al curso 9. Asimismo lo hicieron los otros 2 alumnos clasificados en este curso en una de las sedes externas. En los niveles 11 y 12 no hubo personas inscritas.

A continuación se presenta la tabla que muestra el curso real en que los estudiantes evaluados participaron.

Tabla 10.

Curso a inscribir * Curso en el que participa Crosstabulation

Count		Curso en el que participa											Total	
		Ingles 1	ingles 2	Ingles 3	Ingles 4	Ingles 5	Ingles 6	Ingles 7	Ingles 8	Ingles 9	Toefl	ninguno		
Curso a inscribir	Ingles 1	29											11	40
	Ingles 2	3	4										6	13
	Ingles 3	1		18	1								10	30
	Ingles 4	1			12	1							4	18
	Ingles 5					20							7	27
	Ingles 6						11						13	24
	Ingles 7							20					6	26
	Ingles 8								7				9	16
	Ingles 9									16			4	20
	Ingles 10									3			9	12
	Ingles 11												5	5
	Ingles 12												1	1
	Toefl									1	14		2	17
Total		34	4	18	13	21	11	20	7	20	14	87	249	

Como se mencionó anteriormente, una vez que a los alumnos se les entrevistó e indicó el curso al que debían inscribirse, los candidatos llenaron un cuestionario que tenía por objeto

determinar la evidencia de validez del instrumento escrito y del proceso de clasificación en general.

Se les hizo 3 preguntas fundamentales (con sus correspondientes alternativas en caso de marcar alguna respuesta de manera negativa):

1. ¿Consideras que el proceso de clasificación que acabas de realizar te ubicó en el nivel adecuado?
2. Si repitieras el proceso de clasificación el día de mañana (sin estudiar para llevarlo a cabo) ¿crees que el resultado determinaría el mismo nivel?
3. De acuerdo con el resultado obtenido ¿es tu intención inscribirte al curso asignado?

El resultado general a estas preguntas fue sumamente positivo, como puede verse en las tablas 11 y 12:

Tabla 11. Cuest 1 a) ¿El proceso te ubicó en el nivel adecuado?

		Frequenc	Percen	Valid	Cumulativ Percen
Valid	sí	237	95.2	95.2	95.2
	no	12	4.8	4.8	100.0
	Total	249	100.0	100.0	

Es interesante observar que de todos los candidatos que respondieron esta pregunta, todos confirmaron que sí habían sido clasificados en el nivel correcto. Las razones de que doce alumnos no respondieran el cuestionario fueron atribuidas a que el alumno tenía prisa y dijo que regresaría, o porque las profesoras a cargo olvidaron entregarlo.

Para la segunda pregunta se obtuvo el resultado siguiente:

Tabla 12.

Cuest 1 c) Si repitieras el proceso mañana sin estudiar ¿obtendrías = result?

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid sí	228	91.6	91.6	91.6
no	8	3.2	3.2	94.8
no respondió	13	5.2	5.2	100.0
Total	249	100.0	100.0	

El objetivo de esta pregunta fue confirmar si los estudiantes consideraban que el proceso fuera confiable. A los que respondieron que no, se les pidió establecer por qué y se les ofreció indicar si existió algún factor que no les permitiera dar lo mejor de sí al realizar el proceso (por ejemplo: enfermedad, ruido, presión, etc.). Como se observa, 13 candidatos no respondieron, (incluyendo a 12 que no respondieron el cuestionario completo, y 1 que no respondió esta pregunta en específico), y sólo ocho (el 3.2 %) indicó que no. Las razones que establecieron son las siguientes:

Tabla 13. Razones por las cuales, si repitieran el proceso, no obtendrían el mismo resultado.

NIVEL	CURSO	RAZONES
1	2	Indicó que tenía mucho sin contacto con el idioma y que tal vez si hubiera estudiado tendría una mejor ubicación.
2	3	Indicó que un repaso a sus libros que llevaba en el curso anterior, le ayudaría a recordar algunas cosas.
2	3	No indicó las razones.
2	3	Indicó que no porque el examen escrito le sirvió como repaso.
2	4	Indicó que no porque tiene tiempo que no lo practica (Nota: la pregunta no parece haber quedado clara).
3	5	No indicó las razones.
4	8	No indicó las razones.
4	13 (toefl)	Indicó que no, aunque aclaró que no se debía a alguno de los factores como enfermedad, ruido, presión, etc., sino porque no tuvo tiempo de terminar ya que tuvo que irse a otro compromiso.

En la tercera pregunta se trató de detectar a los candidatos que, una vez terminado el procedimiento, tenían la intención de inscribirse a los cursos. Se obtuvo lo siguiente:

Tabla 14.

Cuest 1 e) De acuerdo con el resultado ¿es tu intención inscribirte al curso?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	sí	225	90.4	90.4	90.4
	no	11	4.4	4.4	94.8
	no respondió	13	5.2	5.2	100.0
	Total	249	100.0	100.0	

Nuevamente observamos que un alto porcentaje de alumnos pretendía inscribirse al curso en el que se le clasificó. Hoy sabemos que sólo lo hizo el 65.1%. Las dos razones más generalizadas que se detectaron de por qué no todos lo hicieron fueron la falta de recursos económicos, y que en el caso de la empresa Bimbo algunos de los empleados no obtuvieron la autorización de sus jefes, o el horario determinado por la empresa no les fue conveniente.

Con respecto a la intención de los candidatos de inscribirse al curso como se observa, doce no contestaron el cuestionario, 1 no respondió esta pregunta, y 11 indicaron que no lo harían. Las razones que estos últimos expresaron fueron las siguientes:

Tabla 15.

Razones para no inscribirse al curso

NIVEL	CURSO	RAZONES
3	6	Indicó que el horario no era adecuado pues se empataba con el horario de clase que tenía en la escuela. Agregó que buscaría otra escuela.
3	6	Indicó que el horario no era adecuado por sus clases de licenciatura.
3	6	Indicó que esperaba inscribirse al curso de preparación para el TOEFL y como no fue así no se inscribiría.
3	6	Indicó que no tenía tiempo.
3	6	Indicó que no estaba viviendo en el D.F. y que no se inscribiría por eso y por su trabajo.
3	6	Indicó que por cuestión de horarios.
4	7	Indicó que necesitaba otro curso y no en el que fue clasificada.
4	7	Indicó que el horario no era adecuado por sus clases de licenciatura.
4	7	No indicó la razón.
5	10	Indicó que se equivocó de examen. Quería hacer el examen TOEFL para ser profesor de inglés.
6	12	Indicó que se inscribiría directamente al TOEFL para aprobar el pre-requisito.

Por lo que se refiere al grado de satisfacción del nivel y curso en el que los candidatos fueron clasificados, se pretendió relacionar el resultado con su propio punto de vista con respecto a su nivel de conocimiento del idioma, pues de acuerdo LeBlanc y Painchaud (1985, en Murray 1999), la auto-evaluación puede ser un factor a considerar en los procesos de clasificación puesto que en muchos casos los alumnos conocen el grado de habilidad que poseen en el idioma. Como se comentó previamente, esta información fue incluida en la página 2 del formato de datos personales que llenaron al iniciar el proceso, donde se les pidió determinar el nivel que consideraban tener con base en el cuadro de clasificación del MERCL incluido en el formato.

La siguiente tabla indica los niveles en los que fueron clasificados los distintos estudiantes:

Tabla 16. Nivel de ubicación

		Frequenc	Percen	Valid	Cumulativ Percen
Valid	Principiante	53	21.3	21.3	21.3
	Elemental	48	19.3	19.3	40.6
	Pre-Intermedio	51	20.5	20.5	61.0
	Intermedio	59	23.7	23.7	84.7
	Intermedio-Alto	32	12.9	12.9	97.6
	Avanzado	6	2.4	2.4	100.0
	Total	249	100.0	100.0	

Como se observa a continuación, un porcentaje alto de estudiantes se ubicó en el mismo nivel determinado por los profesores o en un nivel muy cercano a su punto de vista personal, por lo general más bajo, lo que probablemente fue positivo para que la mayoría de los estudiantes se sintieran correctamente clasificados ya que implicaba poder trabajar en un nivel más alto que el que ellos consideraron.

Tabla 17.

Nivel de ubicación * Nivel de autoevaluación del alumno con base en MECRL Crosstabulation

Count

		Nivel de autoevaluación del alumno con base en MECRL						Total	
		No respondió	Principiante	Elemental	Pre-Intermedio	Inermedio	Intermedio -Alto		Avanzado
Nivel ubicación	Principiant	3	42	7	1			53	
	Element	3	15	21	6	2		48	
	Pre-Intermedio	2	4	10	17	17		51	
	Intermedio	5		1	13	29	9	59	
	Intermedio Alto	3		1	6	10	10	32	
	Avanzado	1				1	3	6	
Total		17	61	40	43	59	22	7	249

Podemos resumir los siguientes resultados:

- En el nivel principiante el 79% se clasificó correctamente, y sólo un 15% consideró estar en un nivel más alto. El 6.6% no respondió
- En el elemental se clasificó bien el 44% y el 31% en un nivel menor. El 19% consideró tener un mejor nivel. No respondió el 6%.
- En el caso del nivel pre-intermedio sólo el 33% se ubicó como lo hicieron las profesoras, el 28% en un nivel menor y otro 35 % en el nivel más alto, lo que indica la posibilidad de que a los alumnos les sea difícil determinar el término medio entre los 2 niveles básicos (Principiante y Elemental), y el Intermedio. El 4% no respondió.
- En el nivel intermedio el 49% se consideró ahí, y sólo un 19% pensó que estaba más arriba. El 24% se clasificó más bajo. No hubo respuesta del 8%.
- Por lo que respecta al nivel intermedio-alto el 31% se ubicó ahí, otro 53% más bajo, y sólo 6% más alto. El 10% de los sujetos no respondió
- En el caso del avanzado, el 17% se clasificó como lo hicieron las profesoras, el 66% más bajo, y no se obtuvo información del 17%.

El número de estudiantes evaluados no son suficientes para corroborar la teoría de LeBlanc y Painchaud (1985, en Murray 1999), pero sirvieron de referencia cuando la evaluación oral y escrita no daban suficientes elementos para determinar el nivel del alumno, situación poco común en el proceso de clasificación pero que se presentó y fue resuelta con esta base.

Por otro lado, una vez iniciados los cursos y como se mencionó anteriormente, a las dos semanas del inicio de clases en la UIA y a las tres semanas de iniciadas en las instalaciones externas, se aplicó un segundo cuestionario para recoger información sobre si el alumno sentía que ya en la práctica había sido bien o mal clasificado.

El resultado se muestra en las tablas 18 y 19:

Tabla 18. Cuest 2 ¿Consideras que el nivel es correcto?

		Frequenc	Percen	Valid	
Valid	dejo de asistir.	12	7.4	7.4	
	sí	130	80.2	80.2	
	no	10	6.2	6.2	
	n/r	10	6.2	6.2	
	Total	162	100.0	100.0	

Como se observa, el 80.2 % consideró que el proceso lo había clasificado en el curso correcto; 7.4 % dejó de asistir antes de aplicar el cuestionario (las razones pueden ser varias y desafortunadamente no pueden reportarse). El 6.2% no respondió (en algunos casos los profesores no les dieron a tiempo el cuestionario o el alumno no lo devolvió). Sólo el 6.2% consideró que no había estado bien clasificado.

En los casos en los que el alumno no se consideró bien clasificado, se le preguntó si consideraba que su nivel era más bajo o más alto. A continuación se muestran los resultados:

Tabla 19. Si no, ¿considera su nivel más alto o más bajo?

		Frecuenc	Percen	Valid	Cumulativ Percen
Valid	adecuad	130	52.2	52.2	52.2
	Más bajo	4	1.6	1.6	53.8
	Más alto	6	2.4	2.4	56.2
	n/r	22	8.8	8.8	65.1
	no	87	34.9	34.9	100.0
	Total	249	100.0	100.0	

La tabla anterior muestra que 4 de los 10 alumnos que respondieron no sentirse bien clasificados consideraron que su nivel era más bajo, y 6 de ellos que era más alto.

En este cuestionario se les pidió comentar si habían expresado su punto de vista al profesor o a la coordinación de idiomas y cuál había sido la respuesta de ambos. Los comentarios de los alumnos que expresaron no estar en el nivel adecuado fueron los siguientes:

Tabla 20. Comentarios de alumnos que consideraron estar mal clasificados.

Nivel	CURSO en el que participa	Se considera	¿Lo comentó?	Comentarios
1	1	Más alto	No	Ninguno
1	1	Más alto	Sí	Antes de aplicarse el cuestionario lo comentó en la Coordinación de Idiomas y con la profesora. Se le invitó a visitar la clase del curso 3 elemental, pero comentó que era demasiado alta. Al aplicarse el cuestionario comentó "Estoy a gusto y además me sirve de repaso".
1	2	Más bajo	No	"Estoy repasando, el profesor me dijo que iba a adquirir más vocabulario".
2	4	Más bajo	Sí, a la profesora que aplicó examen	"Siento que mi nivel era más bajo pero la maestra nos ha ayudado mucho a perder el miedo." También indicó que se sentía satisfecha con el comentario de la profesora de que sí podía con ese nivel porque eso le dio seguridad para ponerse al corriente con sus demás compañeros.
2	4	Más bajo	No	"Me siento a gusto con el grupo, me he puesto a repasar. La maestra me ha apoyado repasando algunos temas junto con todo el grupo".
2	Inicial: 4 A las 2 semanas:5	Más alto	Sí, con profesora y Centro Idiomas	"Cuando estaba en el nivel 4 comenté en el Centro de Idiomas el problema que tenía y me pusieron en el nivel 5. Me siento muy a gusto porque siento que estoy aprendiendo muchas cosas nuevas".
4	7	Más alto	Sí, con su profesora	La maestra del curso le comentó que había temas que debía repasar y se sintió satisfecha con la respuesta.

4	7	Más alto	No	No le pareció importante y prefirió repasar.
5	9	Más bajo	No	“Me estoy esforzando por ponerme al nivel”.
5	9	Más alto	No	“Me siento a gusto con el grupo además de que considero bueno repasar y recordar conocimientos vistos con anterioridad”.

Por lo que respecta a los profesores también se les solicitó llenar un cuestionario para conocer si habían detectado alumnos con un nivel diferente al curso que impartían. Los resultados fueron los siguientes:

Tabla 21. Profesores que detectaron alumnos con diferente nivel al impartido.

NIVEL	CURSO	SEDE	No. de estudiantes que asistían al curso cuando se aplicó el cuest.	ALUMNOS MAL CLASIFICADOS (comentarios de los profesores)
1	1	UIA	10	Uno. “Nivel más alto en producción y comprensión. Visitó la clase del nivel 3 pero prefirió permanecer en éste y trabaja muy bien. No representa problemática.” Nota: Su calificación final fue de 9.
2	3	UIA	11	Ninguno
2	4	UIA	8	Dos. “De uno de ellos su nivel es superior al resto del grupo. Ha tenido que ser muy paciente ya que el ritmo suele adaptarse a la mayoría.” La otra persona fue cambiada al curso 5 <u>Nota:</u> Al responder el cuestionario, el alumno que permaneció en el curso indicó que su nivel sí era adecuado. Terminó el curso con una calificación de 9. La persona que cambió de curso obtuvo una calificación final de 9.
3	5	UIA	14	Ninguno.
3	6	UIA	13	Ninguno.
4	7	UIA	17	Ninguno.
4	8	UIA	10	Ninguno.
4	13 (toefl)	UIA	9	Ninguno.
4	13 (toefl)	UIA	17	Uno. “Su nivel es más bajo en gramática. Se esfuerza con sus tareas pero es muy tímida para participar. Está trabajando en parejas con diferentes alumnas mas avanzadas que le están ayudando a comprender mejor y a desarrollar mejores estrategias”. Nota: Su calificación final fue de 8.
5	9	UIA	10	Ninguno.
1	9	Corp. Bimbo	9	Uno. “Su nivel es más alto. Se desempeña bien y está interesado. No me ocasiona problemas”. <u>Nota:</u> Su calificación final fue 10.
2	4	Corp. Bimbo	5	Dos: Sobre uno de ellos: “Nivel más bajo en producción oral. No ha estado en contacto con el idioma por mucho tiempo. Ha estado revisando estructuras y esto ha mejorado su desempeño en clase. Sobre el segundo “Su problema es producción escrita. Está trabajando fuerte y revisa con sus compañeros las estructuras que ya han visto. Sus compañeros lo apoyan mucho”

				<u>Nota:</u> Anteriormente se reportó a dos estudiantes que tuvieron que participar en este curso por no haberse abierto el que les correspondía. Obtuvieron 8 y 9 como calificación final.
3	5	Corp. Bimbo	3	Ninguno.
5	9	Corp. Bimbo	5	Ninguno.
1	1	TecPk Bimbo	19	Uno. <u>Nota:</u> Se detectó que el alumno no hizo examen de clasificación, simplemente entró al curso directamente.
2	3	TecPk Bimbo	7	Ninguno.
2	4	TecPk Bimbo	10	Ninguno.
3	5	TecPk Bimbo	12	Ninguno.
3	6	TecPk Bimbo	5	Uno. "Su nivel es más bajo y tiene problemas en la producción oral". <u>Nota:</u> Esta alumna se incorporó tarde al curso asignado. Cuando la profesora detectó que su nivel era bajo, se le recomendó asistir al curso inmediato anterior pero no regresó.
4	7	TecPk Bimbo	10	Ninguno.
4	8	TecPk Bimbo	2	Ninguno.
5	9	Bimbo	9	Ninguno.

En resumen podemos confirmar que, sin contar los casos en que los alumnos aceptaron desde un principio trabajar en un nivel inferior o superior por falta de grupo, únicamente hubo una participante, de los 10 que expresaron sentirse mal ubicados, que tuvo que ser reclasificada del curso 4 al 5, lo cual como recordaremos fue verificado por la profesora a cargo quien lo consideró apropiado.

Como se reportó en el cuadro relacionado con el cuestionario aplicado a los profesores, esta alumna obtuvo una calificación de 9 sobre 10. Una vez terminado el periodo de clases se le volvió a contactar para preguntarle sobre sus impresiones en cuanto al proceso de evaluación, cambio de grupo, y resultado.

Se le hicieron las siguientes preguntas:

- ¿A qué atribuyes que en la clasificación haya habido un error, al proceso en sí, o a alguna otra razón (por presión, otro compromiso, enfermedad, etc.)?

- ¿Crees que el curso 5 fue definitivamente el correcto o que pudo haberte faltado hacer el curso 4 para tener un mayor avance o un conocimiento más sólido?
- ¿Estás satisfecha con tu desempeño y resultado del curso que terminaste?
- ¿Tienes algún otro comentario relacionado con el proceso de clasificación y el curso que realizaste?

Al respecto hizo los siguientes comentarios: “me puse muy nerviosa al momento de hacer el examen y pasaba por problemas familiares, me sentí muy bien en el curso 5, me sentía segura y entendía todo muy bien, no sentí que me faltara el curso 4, me siento muy contenta de mi desempeño y me gustó mucho el curso, tan es así que me gustaría inscribirme en el nivel 6 el próximo semestre...”

El resto de los diez alumnos que consideraron no estar bien clasificados, decidió esforzarse para subir su nivel o prefirió repasar, o no fue detectado con un nivel diferente por el profesor a cargo del curso (pues como se observará en la tabla 21, sólo se detectaron a 9 personas con un nivel diferente), por lo que no se presentaron más reclasificaciones. En el caso particular de estos estudiantes, en esa misma tabla se observan los resultados de su desempeño en el curso, los cuales fueron también favorables.

▪ **Reporte cuantitativo de resultados del instrumento escrito**

Tal y como se comentó en el marco contextual, la primera versión de una prueba elaborada para el Centro de Idiomas fue la base para el examen que se utiliza actualmente. Dicha versión nunca fue comprobada estadísticamente para conocer su grado de confiabilidad y validez, por lo que consideramos de suma importancia realizar para esta nueva versión un análisis de la confiabilidad de los reactivos tanto por sección, como en su conjunto a lo largo del examen.

El análisis se basó en las respuestas de 558 candidatos que aplicaron el examen escrito en los periodos de Primavera, Verano y Otoño de 2006, y Primavera de 2007. Lo que el examen pretende medir fundamentalmente es la competencia gramatical, aunque de manera integral se evalúa la comprensión de lectura y el uso de vocabulario. Se encuentra conformado por 180 reactivos con 4 opciones a elegir. Los reactivos están divididos en 6 secciones de 30 preguntas cada una. Cada sección representa uno de los 6 niveles de conocimiento establecidos por el MCERL (Principiante, Elemental, Pre-Intermedio, Intermedio, Intermedio-Alto y Avanzado).

Debido al objetivo del examen (conocer las estructuras que maneja el candidato para poder ubicarlo en alguno de los cursos) y por razones prácticas de manejo de tiempo, los candidatos no requieren terminar el examen. Se les pide completar únicamente hasta donde se sientan confiados, procurando terminar por lo menos la sección que hayan comenzado y no dejarla a la mitad.

El análisis estadístico se llevó a cabo con el uso de SPSS para *Windows*, versión 10.0. Inicialmente, para llegar a la versión actual que se ha aplicado desde 2006, se realizó una serie de pruebas de estadística descriptiva para cada reactivo y sus opciones con el fin de modificar aquellas que no fueran elegidas por los candidatos o las que eran elegidas muy pocas veces, procurando que cada opción tuviera un porcentaje de probabilidad de elección.

Una vez que la versión actual quedó lista, se recolectaron las respuestas de los 558 candidatos en SPSS, y posteriormente se utilizó el programa *Item and Test Análisis Program – Iteman*, versión 2.01 para obtener el análisis de confiabilidad e índice de discriminación de los reactivos por nivel, y de las 180 preguntas en su totalidad.

La tabla 22 muestra la estadística descriptiva del análisis de los 30 reactivos correspondientes a cada nivel. En ella se reportan los 558 casos evaluados.

Tabla 22. Estadística Descriptiva por Nivel

	Principiante	Elemental	Pre-Intermedio	Intermedio	Intermedio Alto	Avanzado
No. de reactivos	30	30	30	30	30	30
No. de casos	558	558	558	558	558	558
Media	19.916	16.794	14.142	11.419	8.409	3.857
Desviación Estándar	6.230	7.842	8.310	8.705	8.121	5.142
Asimetría	-0.583	-0.577	-0.360	0.038	0.423	1.071
Curtosis	-0.362	-0.514	-0.816	-1.197	-1.086	0.096
Mínimo	1.000	0.000	0.000	0.000	0.000	0.000
Máximo	30.000	30.000	30.000	30.000	28.000	23.000
Mediana	21.000	18.000	15.000	12.000	8.000	0.000
ALPHA	0.868	0.917	0.929	0.941	0.943	0.909
Error Estándar	2.262	2.255	2.209	2.117	1.942	1.553
Índice de dificultad (Media= Mean P)	0.664	0.560	0.471	0.381	0.280	0.129
Coefficiente de Correlación Biserial (Media)	0.611	0.698	0.731	0.777	0.822	0.830
Correlación Punto Biserial	0.458	0.544	0.572	0.605	0.609	0.507

A continuación se desglosa lo que cada rubro indica:

- La **media** se refiere al promedio (aritmético) de respuestas correctas obtenido en cada sección. Como puede verse, en el nivel principiante, la puntuación promedio fue de 19.916 respuestas correctas de un máximo de 30; 16.794 para el nivel elemental; 14.142 en el caso de Pre-Intermedio; 11.419 para intermedio; 8.409 para el Intermedio-Alto y únicamente el 3.857 en el Avanzado. Es interesante observar cómo decrece el promedio de aciertos conforme se incrementa el nivel de dificultad
- La **desviación estándar** nos indica el promedio de desviación de las puntuaciones con respecto a la media. Esto significa que en el nivel principiante, hubo sujetos que se desviaron de la media por 6.2 puntos, y pudieron haber obtenido por lo tanto entre 13.71 y 26.1 respuestas correctas (la media fue de 19.91); la desviación estándar para el nivel

Elemental fue de 7.8; para Pre-Intermedio 8.3; 8.7 en Intermedio; 8.1 en Intermedio-Alto, y 5.1 en Avanzado.

- La **asimetría** nos dice en qué lado de una curva de distribución se agrupa la mayoría de los resultados. Si es positiva significa que la mayoría de los valores se encuentran del lado izquierdo de la curva (por debajo de la media). Si es negativa, nos indica que dichos valores se encuentran hacia la derecha, es decir, por encima de la media.

En la tabla observamos que la asimetría es negativa para los primeros tres niveles, Principiante, Elemental y Pre-Intermedio, lo que quiere decir que en cada una de estas secciones de la prueba, hubo más resultados altos (por encima de 19.9, 16.7 y 14.1, respectivamente), que bajos. En contraste, al haberse obtenido una asimetría positiva en los niveles Intermedio, Intermedio-Alto y Avanzado, encontramos que estas secciones fueron consideradas más difíciles porque las puntuaciones se ubicaron por debajo de la media, 11.4, 8.4, y 3.9.

- La **curtosis** representa qué tan plana o elevada es la curva de distribución en cuanto a la concentración de los datos. Si fuera 0 sería una curva normal y los datos estarían agrupados alrededor de la media. Si es positiva significa que es más elevada, es decir, habría un elevado nivel de concentración de datos cercanos a la media; si es negativa, es plana o con un grado bajo de concentración de datos alrededor de los valores centrales.

Podemos ver que de los niveles Principiante a Intermedio-Alto, la curva es plana, lo que significa que en cada nivel hay una alta variabilidad o heterogeneidad en cuanto al conocimiento de los sujetos ya que las puntuaciones están ampliamente distribuidas a lo largo de la escala. En el caso del nivel Avanzado, la curva es elevada, al haber sido positiva, lo que indica que el conocimiento de los alumnos que se encuentran alrededor de la media es más

homogéneo. Cabe mencionar que únicamente 227 alumnos contestaron los primeros 15 reactivos de este nivel, y 202 lo completaron, es decir un 36%.

- Los valores **mínimo** y **máximo** se refieren al mínimo de respuestas correctas (una en el caso de Principiante, es decir que hubo alguien que sólo obtuvo una respuesta correcta, y 0 en los demás casos lo que significa que hubo quienes no tuvieron ninguna respuesta correcta, lo que se debe en parte a que los sujetos no tienen que contestar toda la prueba), y al máximo de aciertos en cada sección, es decir 30 de Principiante a Intermedio, y 28 y 23 en Intermedio-Alto y Avanzado.

- La **mediana** divide la distribución por la mitad y representa el punto medio de la puntuación total. Es decir, en el nivel principiante, el 50% de los sujetos está por encima de 21 respuestas correctas, y el restante 50% por debajo de este valor. En el Elemental la mediana fue de 18; de 15 en el Pre-Intermedio, de 12 en el Intermedio; de 8 en el Intermedio Alto; y para el nivel Avanzado la mitad de quienes lo contestaron no obtuvo ninguna respuesta correcta, y la otra mitad sí.

- El **Alfa** (Alfa de Cronbach) se refiere al grado de fiabilidad de la prueba y es uno de los índices más importantes en este análisis, ya que nos habla de la consistencia interna del instrumento, es decir, del grado de relación que existe entre los reactivos: si estos miden lo que deben medir, y por lo tanto si son homogéneos y están relacionados entre sí. El grado de confiabilidad puede variar entre 0 (que representaría una confiabilidad nula) y 1 (confiabilidad perfecta). McDonald (2002) reporta que el grado de confiabilidad que por lo general se consigue en los exámenes elaborados por profesores tiende a variar entre 0.50 y 0.85, siendo éste el rango mínimo de confiabilidad aceptable cuando se toman decisiones con base en un solo resultado.

Como puede observarse en la tabla, el grado de confiabilidad del nivel Principiante es de 0.87; de 0.92 para el nivel Elemental; 0.93 para Pre-Intermedio; 0.94 para Intermedio; 0.94 para Intermedio-Alto, y 0.91 para Avanzado, lo que habla de un alto grado de homogeneidad entre los reactivos de cada sección del instrumento escrito.

- El **error estándar** es un indicador de la precisión de la media y en la tabla se aprecia una posible variación de 2.26 para el nivel Principiante, 2.25 para el Elemental, 2.21 para el Pre-Intermedio, 1.94 para el Intermedio Alto, y 1.55 para el Avanzado.
- El **índice de dificultad** nos da información sobre la proporción de sujetos que respondieron correctamente los reactivos de cada nivel, por lo tanto indica la dificultad de cada sección. Mientras más alta la proporción, más fáciles fueron los reactivos.

Lafourcade (1973, citado en Chávez, 2001), establece los siguientes criterios para determinar el índice de dificultad:

Tabla 23. Índice de dificultad (Lafourcade, 1973).

ÍNDICE	REACTIVOS
> 85%	Muy fáciles
Entre 50 y 85%	Relativamente fáciles
Entre 15 y 50%	Relativamente difíciles
Entre 0 y 15 %	Muy difíciles

En la tabla 22 podemos apreciar que la sección más fácil de toda la prueba fue Principiante, con 0.66, y que el examen se va haciendo más difícil puesto que el índice disminuye gradualmente (0.56 en Elemental, 0.47 en Pre-Intermedio, 0.38 en Intermedio, 0.28 en Intermedio-Alto) hasta llegar a la sección de mayor dificultad, el nivel Avanzado, con un índice de 0.13. El nivel de dificultad va en aumento como se esperaría que sucediera en un examen cuyo objetivo es establecer cuánto conocimiento tiene el candidato para poder determinar su nivel.

Por su parte, Backhoff, et al (2000) afirman que el nivel medio de dificultad debe oscilar entre 0.5 y 0.6, distribuyéndose los valores de p (dificultad) en 5% de reactivos fáciles, 20% medianamente fáciles, 50% con dificultad media, 20% medianamente difíciles, y 5% difíciles. En la sección de análisis de reactivos compararemos la distribución de dificultad de los reactivos del examen para revisar cómo se comportan con respecto a este criterio.

- El **índice de discriminación** (reportado en las dos últimas filas de la tabla) indica el grado en el que los reactivos distinguen entre los sujetos. Se esperaría que quienes obtuvieron una puntuación alta en todo el examen, tuvieran alta probabilidad de contestar correctamente un reactivo en específico, y viceversa.

Mientras más alto sea el índice de discriminación (en un rango de 0 a 1) se considera que el reactivo distingue mejor a las personas con altas y bajas calificaciones.

Para poder analizar la calidad de los reactivos en este estudio nos basamos en la siguiente tabla de Ebel y Frisbie (1986), en Backhoff, Larrazolo, Rosas, (2000).

Tabla 24. Índice de discriminación (Ebel y Frisbie, 1986)

Discriminación =	Calidad	Recomendaciones
> 0.39	Excelente	Conservar
0.30 – 0.39	Buena	Posibilidades de mejorar
0.20 – 0.29	Regular	Necesidad de revisar
0.00 – 0.20	Pobre	Descartar o revisar a profundidad
< - 0.01	Pésima	Descartar definitivamente

- Ahora bien, el **coeficiente de correlación biserial** determina el grado en que las competencias que mide el test también las mide el reactivo. Nos da un estimado de la correlación entre la calificación total de la prueba y el continuo hipotético del reactivo cuando éste se dicotomiza en respuestas correctas e incorrectas (Backhoff, et al 2000). Observamos en la tabla que el nivel de discriminación a este respecto es muy alto en todos

los casos, pues se incrementa de acuerdo al nivel, siendo en el Principiante de 0.61, y en el avanzado de 0.83.

- De acuerdo con estos autores, la **correlación del punto biserial** se utiliza para saber si las personas “adecuadas” son las que obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo, y cómo puede contribuir a las predicciones. Asimismo, combina la relación entre el criterio del reactivo y el nivel de dificultad. Los autores mencionan que de acuerdo con Ebel y Frisbie (1986) la correlación del punto biserial describe la relación entre las respuestas a un reactivo y las calificaciones en el examen de todas las personas, por lo que es un indicador mucho más preciso que el anterior.

Como puede notarse, el punto biserial de cada nivel es también alto ya que se reporta un 0.46 en el nivel principiante, 0.54 en Elemental; 0.57 en Pre-Intermedio; 0.61 tanto en Intermedio como en Intermedio-Alto, disminuyendo un poco en el caso de Avanzado, puesto que el resultado fue de 0.51. Sin embargo, de acuerdo con la tabla de Ebel y Frisbie (1986 en Backhoff et al 2000) podemos confirmar que en todos los casos se considera como un grado excelente de discriminación. La tabla 25 muestra el resultado concentrado de todos los niveles y los 180 reactivos.

Tabla 25. Estadística Descriptiva (Prueba completa)

	Todos los Niveles
No. de reactivos	180
No. de casos	558
Media	74.536
Desviación Estándar	39.424
Asimetría	-0.038
Curtosis	-0.981
Mínimo	1.000
Máximo	162.000
Mediana	81.000
ALPHA	0.982
Error Estándar	5.259
Índice de dificultad (Media= Mean P)	0.414
Coefficiente de Correlación Biserial (Media)	0.647
Correlación Punto Biserial	0.482

Observamos que el promedio de respuestas correctas fue de 74.5, con una desviación estándar de 39.42, es decir sujetos que obtuvieron esa cantidad de reactivos correctos menos o más con respecto a la media. La asimetría del examen se reporta como negativa, por lo que la curva de distribución se inclina hacia la derecha, lo que indica la existencia de más valores por encima de la media. En virtud de que la curtosis también es negativa, sabemos por lo tanto que la curva es plana y que el conocimiento de los sujetos está ampliamente distribuido.

Los valores mínimos y máximo nos informan que el mínimo de respuestas correctas en todo el examen fue de 1, y el máximo 162. Asimismo, de acuerdo con los datos de la mediana, podemos observar que el 50% de los sujetos estuvo por debajo de 81 reactivos correctos y el resto por encima de ese valor.

Uno de los factores más importantes, el Alfa, reporta un nivel de confiabilidad del 0.982, lo que puede considerarse como muy alto, puesto que el valor máximo sería de 1. Esto significa que el examen es sumamente consistente en todas las áreas, lo que permitió a los sujetos evaluados desempeñarse de la misma forma a lo largo del examen.

Por lo que respecta a los tres últimos indicadores de la tabla, podemos afirmar (con base en Lafourcade 1973, en Chávez 2001), que en su conjunto, es una prueba relativamente difícil, ya que el índice de dificultad del examen es de 0.41. Asimismo, es un instrumento que discrimina bien entre los sujetos, ya que se reporta un coeficiente de correlación biserial (la correlación total de la prueba) de 0.65, y una correlación del punto biserial (la relación entre el criterio de los reactivos y el nivel de dificultad) de 0.46. Recordemos que Ebel y Frisbie (1986) establecen que un índice de discriminación mayor a >0.39 es considerado como excelente.

- Análisis de Reactivos

De acuerdo con la tabla 23 de Lafourcade (1973, en Chavez 2001), los reactivos que se encuentren en un rango de *relativamente fáciles* (entre 0.50 y 0.85) y los *relativamente difíciles* (entre 0.15 y 0.50) pueden considerarse adecuados para formar parte de un examen. Por el contrario, los considerados *muy fáciles* (arriba de 0.85), y los *muy difíciles* (por debajo de 0.15), deben ser analizados para determinar las causas de su extrema facilidad o dificultad, y deberán corregirse o descartarse, según sea el caso. Para tal efecto se reevaluará el índice de dificultad de la prueba si se elimina cada uno de los reactivos que hayan presentado estos índices.

Por su parte, Ebel y Frisbie (1986, en Backhoff, et al 2000) afirman que el índice de discriminación puede considerarse bueno si es arriba de 0.30, y excelente por arriba de 0.39. Índices entre 0.20 y 0.29 deben considerarse como regulares y deben revisarse, y menores a 0.20 deben descartarse. En este análisis aceptaremos los que estén arriba de 0.30 (aún cuando tengan posibilidad de mejorar), y se revisará el índice de discriminación del examen si se eliminan aquellos que estén por debajo de 0.29, con el fin de sugerir su revisión o supresión definitiva.

El análisis de reactivos se llevó a cabo considerando las 180 preguntas de la prueba. No se reporta cada uno por nivel por ser información demasiado extensa, pero el resultado completo puede encontrarse en el apéndice²⁵.

En la tabla 26 aparecen únicamente los reactivos que estén en el límite y los que no hayan cumplido con los criterios de dificultad y de discriminación sugeridos por Lafourcade (1973 en Chávez 2001) y por Ebel y Frisbie (1986 en Backhoff et al 2000). En el caso de que no hayan cumplido sólo con alguno de los dos criterios, se resalta en negritas el índice que sugiere un problema. Asimismo, se reporta el Alfa (la confiabilidad) si se eliminara el reactivo. Para conservarlo el Alfa debe permanecer sin cambio o descender. En caso de que el Alfa se

²⁵ Apéndice 7.

incremento después de quitar un reactivo, éste debe eliminarse (Kim & Shin, 2006). Cabe recordar que los índices obtenidos en la prueba completa fueron de 0.982 para Alfa (confiabilidad); 0.414 (relativamente difícil) para el índice de dificultad, y 0.482 para la Correlación del Punto Biserial (la relación entre el criterio de los reactivos y el nivel de dificultad).

Tabla 26. Análisis de reactivos fuera de los criterios de dificultad y discriminación.

Nivel	Reactivo	Índice de Dificultad 'p'	Índice de Discriminación (Punto Biserial)	Alfa (si se elimina reactivo)	Índice de Dificultad 'p' (si se elimina reactivo)	Índice de Discriminación (Punto Biserial) (si se elimina reactivo)	Sugerencia
Principiante	1	0.688	0.188	0.982	0.413	0.484	Conservar
Principiante	2	0.875	0.225	0.982	0.412	0.484	Conservar
Principiante	3	0.790	0.278	0.982	0.412	0.483	Conservar
Principiante	13	0.466	0.171	0.982	0.414	0.484	Conservar
Principiante	17	0.697	0.235	0.982	0.413	0.484	Conservar
Principiante	18	0.679	0.259	0.982	0.413	0.483	Conservar
Principiante	21	0.523	0.294	0.982	0.413	0.483	Conservar
Elemental	1	0.597	0.234	0.982	0.413	0.484	Conservar
Elemental	8	0.477	0.281	0.982	0.414	0.483	Conservar
Int-Alto	23	0.149	0.442	0.982	0.416	0.482	Conservar
Int-Alto	24	0.136	0.379	0.982	0.416	0.483	Revisar /Conservar
Avanzado	1	0.066	0.320	0.982	0.416	0.483	Revisar /Conservar
Avanzado	3	0.102	0.307	0.982	0.416	0.483	Revisar /Conservar
Avanzado	4	0.145	0.416	0.982	0.416	0.482	Conservar
Avanzado	5	0.124	0.296	0.982	0.416	0.483	Revisar /Conservar
Avanzado	8	0.041	0.145	0.982	0.416	0.484	Revisar /Conservar
Avanzado	9	0.070	0.279	0.982	0.416	0.483	Revisar /Conservar
Avanzado	15	0.106	0.367	0.982	0.416	0.483	Revisar /Conservar
Avanzado	16	0.043	0.071	0.982	0.416	0.484	Revisar /Conservar
Avanzado	17	0.108	0.383	0.982	0.416	0.483	Revisar /Conservar
Avanzado	21	0.057	0.168	0.982	0.416	0.484	Revisar /Conservar
Avanzado	22	0.038	0.210	0.982	0.416	0.484	Revisar /Conservar
Avanzado	23	0.100	0.351	0.982	0.416	0.483	Revisar /Conservar
Avanzado	24	0.061	0.284	0.982	0.416	0.483	Revisar /Conservar
Avanzado	25	0.057	0.212	0.982	0.416	0.484	Revisar /Conservar
Avanzado	27	0.057	0.180	0.982	0.416	0.484	Revisar /Conservar
Avanzado	28	0.120	0.354	0.982	0.416	0.483	Conservar
Avanzado	29	0.034	0.222	0.982	0.416	0.484	Revisar /Conservar
Avanzado	30	0.066	0.264	0.982	0.416	0.483	Revisar /Conservar

Como se observa en la tabla, son 29 los reactivos (el 16%) que fueron analizados por no cumplir ya fuera con el índice de dificultad o con el de discriminación. Once de ellos no cumplieron en ambos casos. No obstante, al eliminarse el reactivo, en ningún caso varió el Alfa.

La dificultad de la prueba se mantuvo como relativamente difícil, ya que el índice varió entre 0.412 y 0.416. El índice de discriminación del punto biserial tampoco reportó ninguna diferencia significativa pues el resultado osciló entre 0.482 y 0.484.

Por lo anterior, en virtud de que no hubo variación en ninguno de los tres elementos analizados, en todos los casos se sugiere conservar el reactivo. Sin embargo, en los casos indicados como *Revisar/Conservar*, la sugerencia se hace porque en ellos los sujetos tendieron a elegir con más frecuencia una alternativa diferente a la respuesta establecida, y el programa *ITEMAN* utilizado para analizar la prueba sugiere en algunos de estos casos revisar la respuesta.

A continuación se realiza una revisión más cercana de estos reactivos y se incluyen otros en que los sujetos eligieron alguna alternativa diferente a la correcta en un porcentaje más alto, aún cuando no hayan sido reportados por el programa *ITEMAN* como problema. Se considera importante incluirlos también con el fin de confirmarle al lector que no se trata de una omisión, sino que a pesar de que el reactivo correcto no obtuvo el porcentaje más alto de respuestas, éste posee un buen grado de dificultad y/o que discrimina adecuadamente.

Tabla 27. Análisis de Reactivos: Elemental 11

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Elemental	11 NO reportado por ITEMAN	0.299	0.479	b Más eligen 'a'	n/r (07.7)	36	6		1			43
					a (35.7)	21	45	41	60	31	1	199
					b (29.9)	4	19	33	64	36	11	167
					c (14.5)	6	13	25	24	10	3	81
					d (12.2)	24	23	11	9	1		68
					Total	91	106	110	158	78	15	588

La alternativa 'a' es elegida por un mayor porcentaje de alumnos de nivel bajo (de Principiante a Pre-Intermedio), mientras que la 'b' la eligen más los de nivel Intermedio, Intermedio-alto y Avanzado. Por tal motivo, el índice de dificultad es de 'relativamente difícil', y el de discriminación se encuentra en el rango de 'excelente'.

El reactivo evalúa el uso de preposiciones de tiempo y lugar.

11. David was born _____ Texas _____ September _____ night.
 a. in ...on ... in the b. in ...in ... at c. on ... in ... at d. in...in...at the

Tabla 28. Análisis de Reactivos: Intermedio 23

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Intermedio	23 NO reportado por ITEMAN	0.246	0.529	b Más eligen 'd'	n/r (31.9)	78	72	26	1	1		178
					a (04.3)	4	6	10	4			24
					b (24.6)	4	5	3	51	49	15	137
					c (01.6)	1	2	2	3	1		9
					d (31.9)	4	21	59	99	27		210
					Total	91	106	110	158	78	15	558

La alternativa 'd' es elegida por el mismo número de sujetos que eligen 'b' y 'a', y que fueron ubicados en Principiante, así como por un gran número de candidatos ubicados del nivel Elemental al Intermedio. La opción correcta, 'b', la eligen más los de nivel Intermedio-Alto y todos los que se ubicaron en Avanzado. Por tal motivo, el índice de dificultad se encuentra en 'relativamente difícil', y el de discriminación se encuentra en el rango de 'excelente'.

La pregunta mide el uso de auxiliares para expresar acuerdo o similitud.

23. I enjoy weddings very much and _____ my girlfriend Sally.
 a. neither b. so does c. nor does d. either

Tabla 29. Análisis de Reactivos: Intermedio-Alto 8

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Intermedio -Alto	8 NO reportado por ITEMAN	0.190	0.614	b Más eligen 'a'	n/r (40.5)	81	87	41	14	3		226
					a (24.9)	3	13	40	65	18		139
					b (19.0)	3	2	15	42	31	13	106
					c (09.9)	1	3	11	23	16	1	55
					d (05.7)	3	1	3	14	10	1	32
					Total	91	106	110	158	78	15	558

La alternativa 'a' es elegida por el mismo número de Principiantes que eligieron 'b' y 'd', así como los ubicados de elemental a Intermedio. La opción correcta, 'b', la eligen más los de nivel Intermedio-Alto y el 86.6% de los que se ubicaron en Avanzado. El índice de dificultad indica un rango de 'relativamente difícil', y el de discriminación es 'excelente'.

En este reactivo se revisa el uso de auxiliares para expresar *'willingness or refusal'* (i.e. utilizado para indicar 'disposición' o cierta 'negación' a hacer algo o a que algo suceda). En esta oración, la opción 'b' se traduciría al español como 'Mi carro no quiere arrancar'. La opción 'a' representa el típico error de transferencia del español

8. Hi Sue! Could you give me a ride? My car _____.

a. doesn't want start b. won't start c. doesn't want to go d. doesn't go

Tabla 30. Análisis de Reactivos: Intermedio-Alto 11

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Intermedio - Alto	11 NO reportado por ITEMAN	0.208	0.774	a Más eligen 'd'	n/r (41.0)	80	88	43	15	3		229
					a (20.8)	2		8	46	46	14	116
					b (03.8)			7	12	2		21
					c (08.6)	2	4	15	25	2		48
					d (25.8)	7	14	37	60	25	1	144
					Total	91	106	110	158	78	15	558

La alternativa 'd' es elegida por un mayor número de candidatos ubicados del nivel Principiante al Intermedio, mientras que la correcta, 'a', la eligen más los de nivel Intermedio-alto y 93.3% de los que se ubicaron en Avanzado. Por tal motivo, el índice de dificultad ('relativamente difícil') es adecuado, y el de discriminación se encuentra en el rango de 'excelente'.

De acuerdo con la Tabla de Especificaciones, se evalúa el uso del participio presente al inicio de oraciones.

11. _____ lunch, we set off on our journey

a. Having finished b. Finished c. Have finished d. When finish

Tabla 31. Análisis de Reactivos: Intermedio-Alto 20

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Intermedio - Alto	20 NO reportado por ITEMAN	0.167	0.444	a Más eligen 'b'	n/r (44.4)	82	90	47	22	7		248
					a (16.7)		3	14	34	28	14	93
					b (16.8)	4	6	20	42	22		94
					c (15.6)	1	1	23	44	18		87
					d (06.5)	4	6	6	16	3	1	36
					Total	91	106	110	158	78	15	558

La alternativa 'b' es elegida por un mayor número de candidatos ubicados del nivel Principiante al Intermedio, mientras que la correcta, 'a', la eligieron más sujetos con nivel Intermedio-Alto y el 93.3 % de los Avanzados. Aunque se considera un reactivo con un alto índice 'relativamente difícil', y se encuentra dentro del límite. El nivel de discriminación es 'excelente'.

El reactivo mide el uso de auxiliares para expresar '*willingness or refusal*' (i.e. utilizado para indicar 'disposición' o cierta 'negación' a hacer algo o a que algo suceda). A diferencia de la pregunta 8 del mismo nivel, ésta está en pasado, por lo que una posible equivalencia en español sería "Los frenos no funcionaban, por lo que llevé el carro al mecánico".

20. The brakes _____, so I took the car to the workshop.
 a. wouldn't work b. not worked c. weren't work d. aren't worked

Tabla 32. Análisis de Reactivos: Intermedio-Alto 23

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Intermedio - Alto	23 NO reportado por ITEMAN	0.149	0.442	d Igual núm. elige 'a'	n/r (45.9)	81	89	50	26	10		256
					a (14.9)	1	9	20	46	7		83
					b (12.2)	4	4	18	29	11	2	68
					c (12.2)	3	3	13	30	16	3	68
					d (14.9)	2	1	9	27	34	10	83
					Total	91	106	110	158	78	15	558

La alternativa 'a' es elegida por un mayor número de candidatos ubicados en los niveles Elemental, Pre-Intermedio e Intermedio, mientras que la correcta, 'd', la eligieron más sujetos con nivel Intermedio-Alto (el 43.5%), y el 66.6 % de los Avanzados. Es un reactivo que se considera 'muy difícil' pues está ligeramente abajo del rango aceptable. Sin embargo, su índice de discriminación está en el rango de 'excelente'.

La pregunta revisa Patrones Verbales, es decir el uso de forma simple, infinitivo o '*ing*' después de ciertos verbos. En el caso de '*regret*', el verbo a seguir requiere '*ing*', por lo que la opción 'd' es correcta.

23. Susan regrets _____ without a license in a rented car.
 a. drive b. to have drive c. be driving d. having driven

Tabla 33. Análisis de Reactivos: Intermedio-Alto 24

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Intermedio - Alto	24 REPOR-TADO por ITEMAN Se sugirió revisar	0.136	0.379	b Más eligen 'c'	n/r (45.2)	82	90	50	22	8		252
					a (06.1)		2	8	22	2		34
					b (13.6)	3	4	7	25	26	11	76
					c (25.1)	5	5	27	64	35	4	140
					d (10.0)	1	5	18	25	7		56
				Total	91	106	110	158	78	15	558	

La opción 'c' es elegida por más sujetos ubicados en los niveles Principiante a Intermedio-Alto (éste con un % de 44.8), mientras que el único nivel en el que los candidatos eligen la respuesta correcta, 'b' es en el Avanzado (73%). El nivel de dificultad es de 'muy difícil', sin embargo el de discriminación se reporta como 'bueno'.

Como se comentó previamente, al eliminar el reactivo el Alfa de la prueba no varió, por lo que se sugirió conservar. Sin embargo se sugiere la revisión, una vez más, de los expertos de la lengua para determinar si hay algún elemento que deba modificarse, ya sea de la base o de las opciones.

La pregunta evalúa el uso de 'be used to + ing' para expresar hábitos o costumbre. El error típico es utilizar el infinitivo en lugar de 'ing' como es el caso de la opción 'c'

24. I _____ around the city by subway. It's something I do every day.
 a. used get b. 'm used to getting c. 'm use to get d. use to getting

Tabla 34. Análisis de Reactivos: Avanzado 1

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	1 REPOR-TADO por ITEMAN Se sugirió revisar	0.066	0.320	d Más eligen 'a'	n/r (54.3)	83	93	60	48	18	1	303
					a (34.9)	2	11	43	91	41	7	195
					b (03.2)	3		5	6	3	1	18
					c (00.9)	2	1	1	1			5
					d (06.6)	1	1	1	12	16	6	37
				Total	91	106	110	158	78	15	558	

La opción 'a' es elegida por la mayoría de los sujetos ubicados en todos los niveles. Únicamente el 40% de los Avanzados escogieron la respuesta correcta ('d'). Por consiguiente, el reactivo es 'muy difícil', sin embargo el índice de discriminación se reporta como 'bueno'.

Cuando se eliminó el reactivo, el Alfa de la prueba no cambió, por lo que se sugirió conservar. No obstante se sugiere su revisión, para que los expertos de la lengua determinen si hay algún elemento que deba modificarse, o si en efecto puede conservarse.

Cabe hacer notar que el reactivo se refiere a una sutileza del lenguaje como son las "expresiones plurales con adjetivos compuestos" cuya regla indica que al utilizar guiones, el sustantivo que forma parte de la expresión adjetiva, no se pluraliza. La opción 'a' representa el error clásico que cometen los hablantes de español, por lo que, en la opinión de la autora de este trabajo, no debería modificarse.

1. Dave bought a _____ picture of two men working in a field.
 a. three-hundred-years old b. three-hundreds-years old c. three-hundred-old years d. three-hundred-year old

Tabla 35. Análisis de Reactivos: Avanzado 3

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	3 REPOR-TADO por ITEMAN Se sugirió revisar	0.102	0.307	b Más eligen 'd'	n/r (56.5)	83	98	64	50	19	1	315
					a (08.8)	2	1	9	21	11	5	49
					b (10.2)	3	1	6	22	20	5	57
					c (08.1)	1	3	14	20	7		45
					d (16.5)	2	3	17	45	21	4	92
					Total	91	106	110	158	78	15	558

Como se observa, la opción 'b' es elegida por tan sólo un sujeto más en los niveles Principiante y Avanzado, y la 'd' la eligen más candidatos ubicados en los niveles Elemental a Intermedio-Alto. Únicamente el 40% de los del nivel Avanzado escogieron la respuesta correcta ('d'). En consecuencia, el nivel de dificultad es de 'muy difícil'. No obstante, el de discriminación puede considerarse 'bueno'.

Cuando se eliminó el reactivo, el Alfa de la prueba no cambió, por lo que se sugirió conservar, aunque es conveniente que los expertos de la lengua determinen si hay algún elemento que deba modificarse o si puede mantenerse.

Cabe mencionar que el reactivo mide el pasado perfecto, y que el distractor 'd' debería ser descartado debido a que la oración es exclamativa, y no interrogativa, por lo que la estructura: Aux+sujeto+verbo, no es adecuada.

3. Alice sat and looked finally at her new house. What a busy day _____!

a. has it been b. it had been c. was it d. has she had

Tabla 36. Análisis de Reactivos: Avanzado 5

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	5 REPOR-TADO por ITEMAN Se sugirió revisar	0.124	0.296	d Más eligen 'a'	n/r (56.3)	83	94	65	51	20	1	314
					a (27.6)	3	5	24	69	44	9	154
					b (02.2)	1	3	3	5			12
					c (01.6)	2	1	2	3	1		9
					d (12.4)	2	3	16	30	13	5	69
					Total	91	106	110	158	78	15	558

La opción 'a' es elegida con mayor frecuencia por todos los candidatos que respondieron el reactivo. Por lo tanto, tenemos un índice de dificultad de 'muy difícil', fuera del rango aceptable. En cuanto al poder de discriminación, se encuentra cerca del rango 'bueno'.

El Alfa tampoco varió al ser omitido de la prueba, por lo que se recomienda su revisión por expertos.

El reactivo evalúa el manejo de la diferencia entre 'like' y 'as', y la opción 'a' representa el error típico cometido por no nativos hablantes de la lengua.

5. He worked on the land _____.

a. like his father did b. as him c. like did his father d. as his father had done

Tabla 37. Análisis de Reactivos: Avanzado 8

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	8 REPOR-TADO por ITEMAN Se sugirió revisar	0.041	0.145	a Más eligen 'c'	n/r (57.9)	83	94	68	53	24	1	323
					a (04.1)	3	1	3	11	4	1	23
					b (11.5)	1	5	12	22	15	9	64
					c (21.1)	2	4	20	56	32	4	118
					d (05.4)	2	2	7	16	3		30
					Total	91	106	110	158	78	15	558

El reactivo tiene un rango de dificultad muy alto debido a que la mayoría de los sujetos eligen la opción incorrecta 'c'. De acuerdo con los criterios para el índice de discriminación, debería descartarse, pero ni el Alfa ni los índices de dificultad y discriminación para la prueba en su totalidad variaron cuando el reactivo fue eliminado, lo que indica que podría conservarse. Con todo, se recomienda revisar para determinar su cambio o permanencia.

De acuerdo con la Tabla de Especificaciones, el reactivo mide *Probabilidad Pasada con Auxiliares Modales*.

8. _____ Peter you met at the party last night. He was ill in bed.
a. It won't have been b. It hasn't been c. He can't be d. He shouldn't be

Tabla 38. Análisis de Reactivos: Avanzado 9

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	9 REPOR-TADO por ITEMAN Se sugirió revisar	0.070	0.279	d Más eligen 'a'	n/r (57.9)	83	95	68	53	23	1	323
					a (21.5)	5	7	24	61	21	2	120
					b (05.0)		1	3	15	6	3	28
					c (08.6)	3	1	8	16	15	5	48
					d (07.0)		2	7	13	13	4	39
					Total	91	106	110	158	78	15	558

Es un reactivo considerado 'muy difícil' con un índice de discriminación 'regular' por lo que se recomienda su revisión a pesar de que ni el Alfa ni los niveles de dificultad y discriminación de la prueba completa reportaron variación al eliminarlo.

El reactivo mide el uso de 'need' como auxiliar para situaciones innecesarias en pasado. La opción 'a' representa el error común que debería descartarse por la falta de la partícula 'to' del infinitivo que se utiliza después de 'need' como verbo. Tal es el caso de la alternativa 'c' en la que se requeriría el infinitivo para que fuera correcta.

Por otra parte, la opción 'b', que es la menos elegida, debería modificarse ya que la estructura es posible para situaciones innecesarias en el presente, y el contexto de la oración permitiría tanto presente como pasado. Se sugiere cambiar a 'needn't to buy', considerando que este auxiliar no acepta infinitivo después de su uso, o establecer el tiempo de la oración en la base.

9. You _____ any butter. We've got lots... but thank you, anyway...
a. didn't need buy b. needn't buy c. didn't need buying d. needn't have bought

Tabla 39. Análisis de Reactivos: Avanzado 15

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	15 NO reportado por ITEMAN	0.106	0.367	a Más eligen 'c'	n/r (59.3)	84	94	69	55	28	1	331
					a (10.6)		1	13	25	17	9	59
					b (09.1)	1	5	7	24	5	3	51
					c (18.3)	4	6	19	49	23	1	102
					d (02.7)	2		2	5	5	1	15
					Total	91	106	110	158	78	15	558

La dificultad del reactivo es alta ('muy difícil') ya que está por debajo del rango aceptable de 0.15. Sin embargo su rango de discriminación es 'bueno', y no se obtuvo variación ni en el Alfa ni en la dificultad y discriminación general al eliminarlo de la prueba, razón por la que se recomienda revisar con el fin de conservarlo de acuerdo con el consenso de expertos.

El reactivo evalúa el mismo tema que el número 9 ('need' como auxiliar). La opción 'c', elegida por la mayoría, es un buen distractor porque representa el error típico.

15. You _____ me up from the station. I would've taken a taxi.
 a. needn't have picked b. needn't to pick c. didn't need pick d. need not picking

Tabla 40. Análisis de Reactivos: Avanzado 16

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	16 REPOR-TADO por ITEMAN Se sugirió revisar	0.043	0.071	b Más eligen 'c'	n/r (61.3)	84	94	72	59	32	1	342
					a (06.5)		2	7	18	6	3	36
					b (04.3)	5		7	9	2	1	24
					c (19.4)	1	7	15	48	27	10	108
					d (08.6)	1	3	9	25	11		48
					Total	91	106	110	158	78	15	558

El reactivo es 'muy difícil', además de que tiene un índice de discriminación considerado 'pobre'. Inclusive la opción correcta es la que menos fue elegida por los evaluados.

Como se reportó previamente, el Alfa no varió al eliminarlo, pero se recomienda el consenso de expertos.

El reactivo revisa el uso del subjuntivo (*that* + oración completa) después de ciertos verbos como '*suggest*'. El error típico es el uso de objeto + infinitivo, representado en 'c', y con menor frecuencia el de la alternativa 'a' (objeto + '*ing*'). La opción 'd' indica un problema de transferencia del lenguaje español ("mi doctor me sugirió que debería.....).

16. My doctor suggested _____ a long break to recover.
 a. me having b. that I have c. me to have d. me I should have

Tabla 41. Análisis de Reactivos: Avanzado 17

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	17 NO reportado por ITEMAN	0.108	0.383	c Más eligen 'a'	n/r (61.5)	84	94	73	59	32	1	342
					a (12.4)	3	10	17	30	8	1	69
					b (09.5)	1		7	25	17	3	53
					c (10.8)	2	1	5	23	18	10	60
					d (05.9)	1	1	7	21	3		33
Total					91	106	110	158	78	15	558	

El grado de dificultad es de 'muy difícil', sin embargo el índice de discriminación es 'bueno', ya que un gran porcentaje de sujetos ubicados en los niveles Intermedio-Alto y Avanzado, eligen la respuesta correcta.

Se revisa el segundo condicional (*if* + pasado / *would*). La opción 'a' (primer condicional) no hace sentido. Se sugiere conservar.

17. If you _____ from my country you _____ understand us better.
 a. come... will b. had come... would c. came... would d. came... will

Tabla 42. Análisis de Reactivos: Avanzado 21

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	21 REPOR-TADO por ITEMAN Se sugirió revisar	0.057	0.168	b Más eligen 'a'	n/r (62.9)	84	95	78	61	32	1	351
					a (19.0)	3	4	11	47	31	10	106
					b (05.7)	1	3	10	14	1	3	32
					c (02.3)	3	1	3	2	3	1	13
					d (10.0)		3	8	34	11		56
Total					91	106	110	158	78	15	558	

El reactivo es 'muy difícil' y el rango de discriminación se reporta como 'pobre' pues la mayoría de los sujetos no eligen la opción correcta. De hecho, antes eligen la alternativa 'd'.

El Alfa y el resto de los elementos analizados para el examen en su totalidad, no presentaron variación después de eliminar el reactivo, por lo que podría mantenerse, una vez confirmado por expertos.

El reactivo evalúa el uso de inversión con palabras negativas (como 'Rarely') al inicio de oración para expresar énfasis, lo que habla de una sutileza del lenguaje de muy alto nivel. El error más común es precisamente la opción 'a', pues utiliza la estructura Sujeto + Verbo, mientras que el problema de la alternativa 'd' es el uso del verbo en pasado, siendo ésta una oración de tiempo presente.

Se recomienda conservar.

21. Rarely _____ a man of such integrity.
 a. you meet b. do you meet c. you do meet d. you met

Tabla 43. Análisis de Reactivos: Avanzado 22

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	22 REPORTADO por ITEMAN Se sugirió revisar	0.038	0.210	c Más eligen 'd'	n/r (63.4)	84	96	79	62	32	1	354
					a (04.3)	3	3	8	10			24
					b (03.0)		1	5	9	2		17
					c (03.8)	1	1	3	7	5	4	21
					d (24.4)	3	5	15	70	9	10	142
					Total	91	106	110	158	78	15	558

El reactivo es 'muy difícil', y su índice de discriminación se considera 'regular'. Como se observa, la mayoría de los sujetos elige la opción 'd', después la 'a', y en tercer lugar la correcta, 'c'.

El Alfa y demás elementos al analizar la prueba completa después de eliminar el reactivo, no presentaron variación, por lo que podría conservarse después de realizar un consenso con los expertos del idioma.

El reactivo evalúa el mismo tema anterior (el uso de inversión con expresiones negativas como '*At no time*' al inicio de oración para expresar énfasis). Las opciones 'a' y 'b' son incorrectas debido a que la oración requiere el uso de Voz Pasiva (be + pasado participio). El problema común es la 'd' pues el distractor utiliza la estructura Sujeto + Verbo, mientras que lo correcto es invertir éstos como en la alternativa 'c'

Se sugiere mantener el reactivo.

22. At no time _____ of the real situation by her doctor.
a. Anna has informed b. Anna informed c. was Anna informed d. Anna was informed

Tabla 44. Análisis de Reactivos: Avanzado 23

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	23 NO reportado por ITEMAN	0.100	0.351	d Más eligen 'a'	n/r (65.1)	84	96	80	67	34	2	363
					a (15.1)	2	4	12	46	19	1	84
					b (03.6)	1	2	4	8	4	1	20
					c (06.3)	3	1	7	15	8	1	35
					d (10.0)	1	3	7	22	13	10	56
					Total	91	106	110	158	78	15	558

La dificultad del reactivo está en el rango de 'muy difícil', pero su índice de discriminación es 'bueno'. Aunque más sujetos ubicados en los niveles Principiante a Intermedio-Alto eligen la opción 'a', el 66.6% de los avanzados escogen la respuesta correcta, 'd'.

Ni el Alfa ni los niveles de dificultad y discriminación sufrieron cambio cuando se omitió el reactivo, por lo que se sugirió conservarse, aunque siempre es conveniente una revisión más cercana de los expertos.

Este es el mismo uso de los 2 anteriores reactivos (inversión con palabras negativas como '*Hardly*' al inicio de oración para expresar énfasis). El distractor 'a' es sin duda el error común debido a la falta de inversión. La alternativa 'c' es imposible porque al ser Pasado Perfecto, se requiere la forma pasada participia del verbo '*meet*', por lo que la 'd' es la respuesta correcta.

Nuestra recomendación es conservar el reactivo.

23. Hardly _____ when they decided to get married.
a. they met b. they did meet c. had they meet d. had they met

Tabla 45 Análisis de Reactivos: Avanzado 34

Nivel	Reactivo	Índice de Díf. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	24 REPOR-TADO por ITEMAN Se sugirió revisar	0.061	0.284	a Más eligen 'b'	n/r (63.8)	84	96	79	63	32	2	356
					a (06.1)	1		6	11	12	4	34
					b (15.4)	3	6	16	46	14	1	86
					c (11.8)	2	4	6	31	15	8	66
					d (02.9)	1		3	7	5		16
					Total	91	106	110	158	78	15	558
<p>El grado de dificultad es una vez más de 'muy difícil', y el nivel de discriminación se considera 'regular'. De hecho, después de la opción 'b', los candidatos eligen la respuesta 'c', y por último la correcta, 'a'. No hubo variación del Alfa o de los niveles de dificultad y discriminación al sacar de la prueba el reactivo, por lo que se sugirió conservar pero revisar.</p> <p>El reactivo mide el uso de 'Wish + past' para expresar deseos en presente o futuro. La opción 'b' es un distractor común por estar en presente.</p> <p>Se recomienda mantener.</p> <p>24. I really wish Hardy _____ me in the house, but he doesn't. a. helped b. helps c. would help d. has helped</p>												

Tabla 46. Análisis de Reactivos: Avanzado 25

Nivel	Reactivo	Índice de Díf. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	25 REPOR-TADO por ITEMAN Se sugirió revisar	0.057	0.212	a Más eligen 'b'	n/r (63.8)	84	95	80	62	33	2	356
					a (05.7)	2	2	4	15	3	6	32
					b (19.4)	1	3	16	44	37	7	108
					c (08.6)	3	3	9	31	2		48
					d (02.5)	1	3	1	6	3		14
					Total	91	106	110	158	78	15	558
<p>El grado de dificultad es de 'muy difícil', y el de discriminación 'regular'. De hecho, después de la opción 'b', los candidatos eligen más la respuesta 'c', y posteriormente la correcta, 'a'. Ni el Alfa ni los niveles de dificultad y discriminación de la prueba total cambiaron al eliminar el reactivo de la prueba, por lo que se sugirió conservar tras su revisión.</p> <p>El reactivo mide el uso de 'It's time + past' para expresar hipótesis de situaciones que ya debieran haber sucedido. La opción 'b' es un buen distractor por estar en presente, pero la respuesta correcta es la 'a' por utilizar el pasado.</p> <p>Se sugiere conservar.</p> <p>25. It's time _____ your degree. a. you got b. you get c. you would get d. you had get</p>												

Tabla 47. Análisis de Reactivos: Avanzado 27

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	27 REPORTADO por ITEMAN Se sugirió revisar	0.057	0.180	b Más eligen 'd'	n/r (64.2)	84	96	81	62	33	2	358
					a (04.8)	4	2	4	11	6		27
					b (05.7)	2	1	5	18	6		32
					c (02.2)	1	1	2	7	1		12
					d (23.1)		6	18	60	32	13	129
					Total	91	106	110	158	78	15	558

El índice de dificultad es de 'muy difícil', y el de discriminación se considera 'pobre', pues la mayoría elige la opción 'd' y ningún candidato considerado avanzado elige la respuesta correcta.

Sin embargo, no hubo variación del Alfa o de los niveles de dificultad y discriminación al omitir el reactivo de la prueba, por lo que se sugirió su revisión para evaluar su posible permanencia.

El reactivo evalúa el uso del subjuntivo ('that' + oración completa) después de adjetivos, en donde el verbo no se conjuga (como en la opción 'd'), sino que se utiliza en forma simple. La opción 'd' representa el error típico.

Se recomienda conservar pues el uso correcto indica un sólido nivel avanzado.

27. It's vital that he _____ his will as soon as possible.
a. would write b. write c. wrote d. writes

Tabla 48. Análisis de Reactivos: Avanzado 29

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	29 REPORTADO por ITEMAN Se sugirió revisar	0.034	0.222	b Más eligen 'c'	n/r (65.1)	84	96	82	64	35	2	363
					a (02.7)	1	1	2	8	2	1	15
					b (03.4)	2		3	4	4	6	19
					c (25.6)	4	7	21	72	34	5	143
					d (03.2)		2	2	10	3	1	18
					Total	91	106	110	158	78	15	558

El reactivo es 'muy difícil', y el índice de discriminación 'regular'.

No hubo cambio en el Alfa o en los niveles de dificultad y discriminación al sustraer la pregunta del examen, por lo que la recomendación fue revisar para determinar su posible permanencia.

El reactivo revisa el uso del segundo condicional con inversión (en lugar de 'if'). El distractor 'c' es incorrecto debido a que requiere el uso de pasado perfecto (*If he had realized*) para ser correcto. Las opciones 'a' y 'd' no son posibles pues su estructura es de oración interrogativa (Aux+Sujeto+Verbo), y la oración no lo es.

29. _____ the danger he was in, he would not proceed with his plan.
a. Was he realizing b. Were he to realize c. If he realize d. Did he realize

Tabla 49. Análisis de Reactivos: Avanzado 30

Nivel	Reactivo	Índice de Dif. 'p'	Índice de Discr. (P Bis.)	Alternativa correcta	Nivel de ubicación de los 558 evaluados y su elección.							
					Alternativas % de elección	Princ.	Elem.	Pre-Inter.	Inter.	Inter-Alto	Avan.	Total
Avanzado	30 REPOR-TADO por ITEMAN Se sugirió revisar	0.066	0.264	d Más eligen 'b'	n/r (63.8)	84	95	78	63	34	2	356
					a (09.5)		2	7	34	10		53
					b (11.5)	1	7	17	31	7	1	64
					c (08.6)	3	2	3	17	16	7	48
					d (06.6)	3		5	13	11	5	37
					Total	91	106	110	158	78	15	558
<p>El índice de dificultad es muy alto, mientras que el de discriminación se encuentra en el rango de 'regular'. Los sujetos eligen más comúnmente la opción 'b', después la 'c', y por último la alternativa correcta, 'd'</p> <p>No hubo diferencia en el Alfa o en los niveles de dificultad y discriminación cuando se omitió el reactivo de la prueba, por lo que se sugirió revisar para ver la posibilidad de conservar.</p> <p>Aquí se evalúan patrones verbales (el uso de forma simple, infinitivo o 'ing' después de ciertos verbos). El uso de infinitivo (opción 'b') representa un error muy común. El problema con la alternativa 'c' es la falta del pasado participio del verbo 'stole' (i.e. 'stolen'), pues el uso de 'having' sería correcto. El patrón adecuado después del verbo 'deny' es el uso del siguiente verbo con 'ing', por lo que la opción 'd' es la respuesta correcta.</p> <p>La recomendación es conservar la pregunta.</p> <p>30. She denied _____ the money. a. to have stole b. to stole c. having stole d. stealing</p>												

Una vez analizados los casos reportados por *ITEMAN*, y aquellos en que los evaluados eligieron con más frecuencia una opción incorrecta, consideramos que es posible que esto obedezca a que los niveles altos requieren un conocimiento más minucioso del idioma, (lo que pocos de los evaluados poseen), y no a que el reactivo o las alternativas de respuesta estén mal planteadas.

Por otra parte, basados en la distribución del índice de dificultad ideal propuesto por Backhoff, et al (2000), a continuación reportamos los porcentajes arrojados por los reactivos de la prueba aplicada en el Centro de Idiomas.

Como se muestra en la tabla 50, sugerimos una pequeña variación en cuanto al rango de ‘dificultad media’, el cual en nuestra opinión, debe oscilar entre 0.40 y 0.60 para que la distribución sea equilibrada²⁶.

Tabla 50. Porcentaje de distribución del índice de dificultad (real e ideal) de la prueba

% IDEAL	REACTIVOS	Distribución IDEAL del Rango de Dificultad	Distribución IDEAL de los 180 Reactivos	Distribución REAL de Dificultad de los reactivos
5%	Fáciles	0.851 a 1.000	9	1 % = (1)
20%	Medianamente fáciles	0.601 a 0.850	36	20 % = (36)
50%	Dificultad media	0.400 a 0.600	90	29 % = (52)
20%	Medianamente difíciles	0.150 a 0.399	36	39 % = (71)
5%	Difíciles	0.000 a 0.149	9	11 % = (20)

El porcentaje más alto de distribución (39%) se ubica en los reactivos ‘medianamente difíciles’ (71). Sólo el 1% es ‘fácil’ (uno de ellos); el 20% (36 preguntas) son ‘medianamente fáciles’ como se espera de manera ideal; el 29% (52 casos) tienen una ‘dificultad media’; y el 11% (20 ítems) son ‘difíciles’.

²⁶ Los autores mencionados sugieren (como se muestra en la tabla 23), que el nivel medio de dificultad debe oscilar entre 0.5 y 0.6, distribuyéndose los valores de p (dificultad) en 5% de reactivos fáciles, 20% medianamente fáciles, 50% con dificultad media, 20% medianamente difíciles, y 5% difíciles.

VI. DISCUSIÓN

▪ Implicaciones

Son dos las implicaciones que pueden surgir del presente estudio. La primera de manera específica para el Centro de Idiomas de la Universidad Iberoamericana sobre el estatus del proceso de clasificación que se utiliza actualmente (lo que se comentará a lo largo de las conclusiones y se extenderá en la sección de recomendaciones). La segunda y más general, para los interesados en el tema de evaluación de idiomas, en particular en los procesos y exámenes de colocación, como son investigadores, administradores de centros de lenguas, y profesores de inglés u otros idiomas.

Importante en cualquier caso es resaltar la trascendencia y transformación que el término *Evaluación* ostenta hoy en día a nivel nacional e internacional. *Evaluar* ha dejado de ser sinónimo de “aprobar o reprobar” y no se limita a un examen o a una calificación.

Por el contrario, *Evaluar* constituye una serie de actividades, tareas, contextos, situaciones, y análisis detallados en torno al desempeño de un individuo con respecto a una competencia en particular, o en cuanto a la eficacia y pertinencia de un programa o institución. El fin último de todo ello conducente a tomar decisiones justas y adecuadas que permitan mejorar, o simplemente determinar, informar, o reconocer un grado de dominio o eficacia.

En el caso de México, todavía encontramos profesores que únicamente toman en cuenta una calificación final sin importar el desempeño, logros y avances del estudiante, sino que se limitan a preguntarse si el alumno “*pasa*” o no.

Asimismo en nuestro país, a pesar de que hay excelentes y dedicados profesores de lengua con amplia experiencia y siempre dispuestos a actualizarse en su campo, son pocas las instituciones de formación de profesores cuyos programas ofrezcan un estudio profundo,

consciente y responsable de las consecuencias de la evaluación, y que a su vez hagan relevantes los procedimientos necesarios para que cualquier proceso de evaluación o instrumento, cumpla por lo menos con los criterios fundamentales de *confiabilidad* y *validez*. La realidad es que muchos maestros de idiomas se enfrentan día a día a la creación y aplicación de exámenes, pero no siempre cuentan con los conocimientos necesarios para elaborarlos con una base científica que asegure que su manera de evaluar sea consistente y mida realmente lo que se pretenda medir.

En nuestro afán de ser mejores cada día como profesores, debemos preguntarnos qué tanto conocemos sobre evaluación y qué tanto aplicamos el nuevo concepto, o si continuamos limitándonos a lo que nos indica un número final.

Es fundamental que quienes se dedican a la enseñanza continúen fomentando o, en su caso, hagan suyo y apliquen gradualmente para su propia práctica, el concepto e implicación de *Evaluar*.

▪ Conclusiones

De acuerdo con los resultados obtenidos, en seguida damos respuesta a las preguntas de investigación que se plantearon al inicio del documento.

Iniciamos con las preguntas específicas que se establecieron con el fin de responder el planteamiento general.

- En un semestre de instrucción ¿existen alumnos que tengan que ser reubicados en niveles distintos al inicial?

En el semestre de Primavera 2007 se inscribieron a los cursos 162 personas de los 249 evaluados (el 65.1%). De los inscritos, el 4% consideró que no había estado bien clasificado. Cuatro de ellos (el 1.6%) indicó que consideraba tener un nivel más bajo, y 6 (el 2.4%) que era más alto. Sin embargo, de estos alumnos se reubicó a una sola

estudiante (con base en la confirmación de los profesores), puesto que el resto, no considero importante solicitar un cambio, o prefirió repasar, o se sintió a gusto en el curso, o el profesor consideró que el cambio no aplicaba y que el alumno estaba en el curso y nivel correctos.

- ¿En cuanto tiempo se determina que el estudiante ha sido mal clasificado y quién se percata de ello (el estudiante, el profesor, la coordinación)?

No se puede determinar cuánto tiempo pasa para darse cuenta de ello, pues las situaciones pueden ser variables.

La mayoría de los estudiantes que consideraron estar mal clasificados no lo expresaron sino hasta que contestaron el cuestionario aplicado dos semanas después del inicio de cursos, e incluso dos de ellos lo consideraron de poca importancia.

La tabla 20 muestra que una persona inscrita en el curso 1 consideró tener un nivel más alto desde el proceso de colocación. Sin embargo, una vez iniciado el semestre visitó el siguiente curso abierto (el curso 3), pero le pareció muy difícil. Dos alumnos que consideraron su nivel más bajo fueron alentados por sus profesoras en cuanto a que podían trabajar en el curso en el que participaban y quedaron satisfechos con el resultado y su desempeño. La persona que fue reubicada detectó estar mal clasificada durante la primera semana de clase, por lo que para cuando respondió el cuestionario, ya había sido cambiada de curso.

El profesor de la persona reubicada detectó el problema los primeros días de clase y confirmó la necesidad del cambio, lo que expuso a la coordinación. Posteriormente, la profesora que recibió a la alumna en su clase indicó que su nivel era adecuado para trabajar en ese programa.

Cabe mencionar que la percepción de los estudiantes no necesariamente coincide con la de sus profesores. De los dos alumnos que expresaron estar más altos, sólo uno fue detectado por el profesor. Del mismo modo, aunque un profesor detectó a un alumno con nivel más alto, el estudiante no lo manifestó. En el caso de una alumna con nivel menor, ella no lo consideró así, aunque el profesor dio a conocer el problema.

En un caso fuera de lo común, el profesor se percató del nivel bajo del alumno, pero al darle seguimiento se detectó que no realizó el proceso de clasificación, simplemente se presentó al curso (lo cual pudo haber ocurrido por estar en sede externa).

Es difícil que las circunstancias se presten para que sea la coordinación la que detecte que el estudiante fue mal clasificado, sin embargo la labor implica estar al pendiente y escuchar y atender los casos manifestados por profesores y alumnos.

Es importante destacar que la aplicación de los cuestionarios después de iniciados los cursos permitió monitorear a los alumnos y consultar a los profesores para, en caso necesario, realizar los cambios pertinentes.

- ¿Cuál es la percepción que tienen los usuarios (alumnos, examinadores y profesores) del proceso de colocación en su conjunto con respecto a su utilidad, confiabilidad y validez?

Los candidatos que realizaron el proceso lo consideraron útil y confiable pues el 95.2% expresó haber sido bien clasificado, lo que nos confirma evidencia de validez.

A continuación se incluyen algunos de los comentarios que hicieron al respecto (se eligió al azar un ejemplo por curso):

Princip 1	“Me pareció bien, es práctico y adecuado. Gracias.”
Princip 2	“Me parece muy interesante que se aplique de esta manera, primero escrita y luego verbal, aunque lo traiciona a uno los nervios al momento de la entrevista en inglés, pero muy bien el proceso.”
Elemental 3	“La evaluación es muy buena y la maestra precisa y objetiva. Es excelente.”
Elemental 4	“Se me hace medio pesado porque son muchas preguntas y hay un momento en que te bloqueas”
Pre-Inter 5	“El examen escrito me gustó y fue fácil de entender. La entrevista fue muy buena”
Pre-Inter 6	“Es un proceso pesado con tantas preguntas pero está bien.”
Intermedio 7	“Todo estuvo bien, es una manera adecuada de hacer una clasificación de nivel”
Intermedio 8	“Muy buena la evaluación, sólo que es un poco extensa. Sugiero hacerla un poco más corta”
Inter-Alto 9	“Sólo dos observaciones: 1) El sistema de llenado de respuestas es muy pesado. Podrían hacer uso de los equipos de cómputo para efectuar los exámenes y así acortar el tiempo, tanto de captura como de espera para las entrevistas. 2) La ubicación del salón no me gustó. El hecho de que esté a un costado de la cafetería aumenta el ruido y el aroma a cigarrillo es desagradable.”
Inter-Alto 10	“Muy adecuado y objetivo el examen de ubicación escrito y oral.”
Avanzado 11	“Pienso que es un buen sistema.”
Avanzado 12	Nota: El único candidato clasificado en este curso no hizo comentarios.
TOEFL	“Muy completa la clasificación, incluye todos los aspectos, aunque es tardado el proceso y a una hora no conveniente por el trabajo.”

Para los examinadores es un proceso largo y cansado principalmente durante los últimos días puesto que la mayor parte de la gente asiste a realizar el proceso hasta los últimos días que se programan para examen.

Otro factor al respecto fue que en esta ocasión el tiempo del proceso se alargó por la necesidad de aplicar los cuestionarios de salida, aunque se procuró que esto se hiciera de manera práctica.

La problemática principal para los examinadores se presenta cuando se acumulan varias entrevistas. En ocasiones se solicita ayuda a otros profesores para calificar el examen escrito (aunque no todos ellos pueden determinar el curso correspondiente por falta de experiencia para hacerlo) para que los evaluadores experimentados puedan realizar la evaluación oral y determinar el nivel y el curso correspondiente. El apoyo de la coordinación está siempre presente en caso necesario.

Por su parte, los profesores consideran que desde hace algunos periodos el nivel de los alumnos es más homogéneo y por lo general cuando llegan a encontrar a alguien con un

nivel diferente es debido a que el curso que le correspondía no se abrió y se le permitió participar en un curso más bajo (o en ocasiones más alto, con esfuerzo). En este periodo fueron nueve los alumnos que, de acuerdo con el criterio de los maestros, presentaron problemas de nivel (el 5.5%), lo que habla de un proceso confiable y útil.

- ¿Existe algún elemento del proceso (examen, entrevista, examinadores, etc.) que pudiera contribuir a la existencia de errores en la clasificación?

Por lo que pudo observarse durante el proceso y con base en los cuestionarios aplicados tanto al terminarlo como una vez iniciados los cursos, se puede confirmar que en su conjunto tanto la aplicación de la prueba escrita, la sección de la entrevista oral y la disposición y buen desempeño de los examinadores se conjugaron adecuadamente para que los errores fueran mínimos.

A pesar de ello, es necesario enfatizar que la sección oral del proceso puede llegar a presentar inconsistencias puesto que existen discrepancias en la forma de evaluar de cada entrevistador. Por ejemplo, el tipo de tarea puede ser muy distinto de candidato en candidato, pues algunos contestan preguntas, mientras que a otros se les pide hablar sobre un tema (su trabajo, o sus estudios, o un viaje, un familiar, etc); y a otros más se les solicita traducir oraciones al inglés. Asimismo, los evaluadores a veces hacen uso exclusivo del inglés y en ocasiones mezclan el inglés y el español. Por otra parte, el tiempo que tardan con los estudiantes puede ser muy variable, a pesar de que en algunos casos los evaluados sustentan niveles similares.

Es muy probable que las inconsistencias mencionadas se deban a que los examinadores que participaron se basan tan sólo en su experiencia (que afortunadamente es amplia) para conducir las entrevistas.

El Centro de Idiomas no cuenta con capacitación específica para entrenar a los profesores en la forma de llevar a cabo una evaluación de esta índole que asegure que se apliquen en todos los casos y de manera uniforme los criterios que les permitan determinar con mayor precisión la competencia oral del alumno, y que eviten que su propio papel como evaluadores influya en el resultado que otorgan a un candidato.

De hecho, algunos profesores del propio Centro prefieren no participar porque consideran difícil realizar el proceso de evaluación para determinar los niveles. En periodos de evaluación anteriores, cuando se le pidió a algún profesor con poca experiencia su apoyo para participar en el proceso, se llegaron a presentar consecuentes errores de colocación.

Respuesta a la pregunta genérica:

- ¿Cuál es la efectividad del proceso de colocación que se lleva a cabo en el Centro de Idiomas de la UIA en cuanto a su confiabilidad y evidencia de validez?

La evaluación de todos los elementos que conforman el proceso de colocación muestra en general resultados positivos que permiten que éste pueda considerarse confiable y que muestre evidencia de validez.

El alto grado de consistencia y estabilidad de la sección escrita permite discriminar de manera efectiva a los candidatos con mayor o menor conocimiento. El reporte estadístico arrojó resultados más que deseables en cuanto al grado de confiabilidad (de 0.98), puesto que se pensaba que pudieran presentarse mayores inconsistencias al respecto.

Asimismo, encontramos sólida evidencia de validez confirmada por expertos del idioma quienes evaluaron de cerca tanto la validez aparente de la prueba (su formato), en cuanto a

que los reactivos y sus opciones fueran claras y relevantes, como la validez de contenido y el uso de la lengua en el examen.

De igual modo, existe evidencia en cuanto a la estructura interna ya que los componentes de la prueba concuerdan entre sí en cuanto al constructo gramatical establecido para evaluar a los candidatos.

Finalmente, el grado de satisfacción reportado por los evaluados, permite sustentar evidencia con respecto a las consecuencias favorables de la evaluación, al haber ubicado adecuadamente al 95.2 % de los candidatos (de acuerdo con su propia percepción), y al 96 % de los inscritos que respondieron el cuestionario aplicado una vez iniciados los cursos.

Por lo anteriormente expuesto, podemos concluir que el proceso de colocación llevado a cabo por el Centro de Idiomas de la Universidad Iberoamericana para determinar el nivel de conocimiento de los evaluados y ubicarlos en el curso de instrucción adecuado a sus necesidades y competencias, es un proceso efectivo, confiable, y que muestra evidencia de validez, elementos que sustentan su efectividad y calidad, como corresponde a la Institución de Educación Superior en la que se ha realizado este estudio.

▪ **Recomendaciones**

Recordemos que uno de los principios de la calidad implica un proceso de mejora continua y con el fin de asegurar que el proceso de colocación mantenga su evidencia de validez y siga siendo confiable, y efectivo, se sugiere:

- Asegurar que la información sobre el proceso en general y las instrucciones para completar la prueba escrita siempre sean claras en cuanto al tiempo y etapas que conlleva.

- Que los expertos de la lengua del Centro de Idiomas lleven a cabo la revisión de los reactivos reportados con inconsistencias (tablas 26 a 49) con el fin de determinar si pueden conservarse o si es recomendable eliminarlos o realizar modificaciones a la base o las opciones. Cabe mencionar que el ítem Avanzado 9, debe ser corregido en cuanto la alternativa *b*) porque como se encuentra actualmente permite que haya dos respuestas correctas (se hace la sugerencia de corrección en el análisis de la tabla 38, p. 94)

En la opinión de quien escribe, el resto de los reactivos reportados debería conservarse sin cambios, ya que las opciones representan los errores típicos incurridos por hablantes de español que aún no manejan esas sutilezas del lenguaje, lo cual permite que la discriminación entre estudiantes sea adecuada.

- Que el Centro de Idiomas con el apoyo de expertos en evaluación, de los profesores de lengua inglesa y otros expertos de la lengua, establezcan con claridad los criterios y el procedimiento para conducir las entrevistas orales, para con esa base impartir la capacitación necesaria, de tal manera que:
 - la evaluación oral sea más consistente y por ende, más confiable,
 - su evidencia de validez sea más sólida,
 - los entrevistadores evalúen a todos los candidatos de manera más justa, con las mismas bases y elementos y para que no incurran (inconscientemente) en aplicar sus propios juicios o juicios ajenos al objetivo de evaluación,
 - todos los examinadores interesados en participar, y los profesores potenciales para apoyar el proceso en un futuro bajo determinadas circunstancias, tengan posibilidad de establecer de manera precisa y confiable el nivel de inglés de un

candidato y el curso más adecuado para él.

Cabe mencionar que Alderson et al (1998) presentan algunas ideas claras y concisas para la capacitación de evaluadores de la competencia oral.

- Hacer uso de los cuestionarios para alumnos y profesores a las dos semanas de iniciados los cursos con objeto de monitorear a posibles estudiantes mal colocados y tomar las decisiones de cambios o soluciones de manera oportuna.
- Seguir utilizando la hoja de información del aspirante con el fin de contar durante la entrevista con mayor información sobre sus antecedentes en el idioma y, en caso necesario, considerar su propia autoevaluación como referencia.

Adicionalmente, esta información puede dar pie a otras investigaciones como si existe correlación entre el nivel de inglés de los estudiantes provenientes de escuelas particulares y el grado de conocimientos de los de escuelas públicas; o si el género, la edad o el grado de estudios es significativo para obtener un mejor desempeño en el curso asignado, etc. Es necesario fomentar mayor investigación en el campo de la evaluación de lengua extranjera para el contexto mexicano.

- Continuar aplicando el cuestionario de salida porque permite monitorear la opinión de los candidatos sobre el proceso. Esto con el fin de asegurar que el nivel de aceptación de la prueba no disminuya o se vea modificado y, en caso de que eso ocurriera, poder actuar a tiempo y tener control sobre cualquier cambio o elemento que provoque una percepción negativa por parte de los aspirantes, e incluso de los evaluadores. De igual modo, este cuestionario permite conocer si un candidato pretende inscribirse o las razones de no hacerlo, y es un medio para que los usuarios hagan sugerencias y aporten de algún modo a la mejora del proceso.

- Que en la medida de lo posible, se cree un banco de ítems que permita modificar periódicamente la prueba escrita y que nuevas versiones sean analizadas estadísticamente para confirmar que discriminen adecuadamente y que sustenten un buen grado de confiabilidad y muestren evidencia de validez de contenido y de estructura interna.
- Que dentro de los proyectos futuros del Centro de Idiomas y de la Institución, la prueba escrita se aplique y califique por computadora pues hacerlo permitiría agilizar el proceso.

De tal modo, los aspirantes podrían realizar el examen en cualquier momento a través de Internet y por el mismo medio solicitarían fecha y hora para su entrevista oral (siempre limitadas a las necesidades prácticas del Centro y a los profesores disponibles para hacerlas, cuyos horarios de entrevista pueden establecerse previamente).

Así, el examen se recibiría y calificaría previamente, lo que pudiera permitir ahorro de tiempo y aumentar la validez aparente del proceso, ya que sería más práctico tanto para los candidatos como para los evaluadores. Más allá de esto, aún cuando la inversión inicial para sistematizar el proceso implique un costo, seguramente éste se vería reflejado en el beneficio que incluiría que los evaluadores inviertan menos tiempo en el proceso y que la Institución reduzca el costo por aplicación de exámenes de colocación.

Estas son tan sólo algunas recomendaciones. Seguramente los interesados podrán aportar valiosas ideas que permitan que el proceso de colocación evaluado aquí mantenga el estatus que se concluye en el presente estudio.

▪ **Comentario final**

La potencial aportación de esta tesis puede resumirse en las siguientes palabras:

Hay mucho que aprender sobre evaluación. Ampliemos nuestro conocimiento, nuestra visión, nuestro interés al respecto. Mejoremos nuestras técnicas para elaborar y aplicar evaluaciones y así poder ser cada día mejores profesores. Nuestros alumnos cuentan con ello.

BIBLIOGRAFÍA

- ALBATCHE, Philip.
“Globalization and the University: Myths and Realities in an Unequal World” en *Tertiary Education and Management* No. 1, 2004.

- ALCARAZ, Enrique y María Antonieta Martínez
Diccionario de lingüística moderna. Barcelona, España, Ed. Ariel, 1997.

- ALDERSON, Charles, Caroline Claphaman, y Diane Wall
Exámenes de idiomas. Elaboración y Evaluación. Madrid, CUP, 1998.

- ANULA Rebollo, Alberto
El abecé de la psicolingüística, Madrid, Arco/Libros, S.L., 1988.

- ARDESHIR, Geranpayeh
“A Quick Review of the English Quick Placement Test” Extract from *Research Notes*,
Vol. 12, 2003.

- BACHMAN, Lyle
Fundamental Considerations in Language Testing. Oxford, Oxford University Press,
1990.

- BACHMAN, Lyle, y Adrian Palmer
Language Testing in Practice. Oxford, Oxford University Press, 1997

- BACKHOFF, Eduardo, Norma Larrazolo, y Martín Rosas
“Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA), en *Redie (Revista Electrónica de Investigación Educativa)*, Vol. 2, Núm. 1, 2000.
<http://redie.uabc.mx/vol2no1/imprimir-contenido-backhoff.html> (consulta Abril 2008).

- BARRÓN Francisco; Sarah Brown, M Cisneros.
Proyecto: Centro de Lenguas. México: ITESO, Dirección General Académica. Depto. de Estudios Socioculturales. Centro de Lenguas, 2002.

- BERNHARDT, Elizabeth, Raymond Rivera, y Mihael Kamil
“The Practicality and Efficiency of Web-Based Placement Testing for College-Level Language Programs”. *Foreign Language Annals 37,3. ProQuest Educational Journals*, 2004.

- BRADSHAW, Jenny
“Test-takers’ reactions to a placement test” en *Language Testing*, Vol. 7, No. 1, Edward Arnold, 1990.

- BROWN, Annie
“Interviewer variation and the co-construction of speaking proficiency” en *Language Testing*, Vol. 20, No. 1, 2003.

- BROWN, James D.
“Improving ESL placement tests using two perspectives”. *TESOL Quarterly*, 23, 1989.

- BROWN, James D.
The elements of language curriculum: A systematic approach to program development.
New York, Heinle & Heinle, 1995.

- CASTILLO Arredondo Santiago
Compromisos de la Evaluación Educativa. Madrid, Pearson Education, 2002.

- CASTILLO Arredondo Santiago
Vocabulario de Evaluación Educativa. Madrid, Pearson Education, 2003.

- CHAVEZ, Luz María
“Análisis de reactivos: una aproximación” en *Biblioteca Universitaria, Nueva Época*,
Vol. 4, No. 2, 2001.
http://www.dgbiblio.unam.mx/servicios/dgb/publicdgb/bole/fulltext/volIV22001/pgs_101-107.pdf
(consulta Mayo 2008).

- CHO, Yeonsuk
“Advanced Placement International English Language (APIEL)” *Language Testing*, vol.
19, No. 3, 2002.

- *Diccionario de la Lengua Española*. Madrid, Real Academia Española, Vigésima Primera Edición, 2000.

- DA SILVA Helena, SIGNORET, Aline.

Temas sobre la adquisición de una segunda lengua. México, UNAM, 2005.

- DAVIES, Alan

Principles of Language Testing. Oxford, Basil Blackwell, Ltd., 1990.

- EBEL, Robert, y David Frisbie

Essentials of Educational Measurement. New Jersey, Prentice Hall, 1991.

- ELOLA, Nydia, y Lilia Toranzos

Evaluación Educativa: una aproximación conceptual. Buenos Aires, OEI (Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura), 2000.

<http://www.oei.es/calidad2/luis2.pdf> (Consulta Agosto 2007).

- *Foreign Language Framework for California Public Schools.* Sacramento: California Department of Education, 2003.

- FULCHER, Glenn

“An English language placement test: issues in reliability and validity” en *Language Testing*, U.K., Arrowsmith, Ltd. Vol. 14, No. 2, Julio 1997.

- GACEL-AVILA, Jocelyne

La internacionalización de las Universidades Mexicanas. Políticas y estrategias institucionales. México: ANUIES, 1998.

<http://www.anuies.mx/principal/servicios/publicaciones/libros/lib43/> (Consulta febrero 2005)

- GACEL-AVILA, Jocelyne

“La dimensión internacional de las universidades mexicanas” en *Revista de la Educación Superior en Línea*, No. 115, Julio 2000.

<http://www.anuies.mx/index1024.html> (Consulta febrero 2005).

- GARCÍA, Sergio

“La validez y confiabilidad en la evaluación del aprendizaje desde la perspectiva hermenéutica” en *Revista de Pedagogía*, Vol. 23, Núm 67, 2002.

http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S0798-97922002000200006&lng=pt&nrm=iso (Consulta: Julio 2008).

- GOODWIN, Laura

“The meaning of validity (news and views)” en *Journal of Pediatric Gastroenterology and Nutrition*, Vol. 35 (1) Julio 2002.

<http://www.jpagn.org/pt/re/jpagn/fulltext.00005176-200207000-00003.htm;jsessionid=L96GKSGnfnxBD0RLQS2zGtLzD1XsQX1sWrgRGqpSYr81y4Tr1RL!536197444!181195628!8091!-1> (Consulta: Julio 2008).

- GOODWIN, Laura, y LEECH, Nancy
“The meaning of validity in the new standards for educational and psychological testing: implications for measurement courses.(Assessment In Action)”. En *Measurement and Evaluation in Counseling and Development*, Oct, 2003. Adapted with permission from *Changing Conceptions of Measurement Validity: An Update on the New Standards*," by L.D Goodwin, 2002 in Journal of Nursing Education. Copyright 2002 by SLACK, Inc.
http://goliath.ecnext.com/coms2/gi_0199-3876917/The-meaning-of-validity-in.html
(Consulta: Julio 2008).

- GREEN, Anthony y Cyril Weir
“Can placement tests inform instructional decisions?” en *Language Testing* Vol. 21, No. 4, 2004.

- HEINKELMAN, Don, y Timothy Grose
“Placement Testing and Audio-Quiz Making with Open Source Software. *PacCall Journall, (Journal of the Pacific Association for the Computer Assisted Language Learning)*, Vol 1, No. 1, 2005.

- HENNING, Grant
A Guide to Language Testing: Development, Evaluation, Research. Boston, Mass., Heinle & Heinle, 1987.

- HERNÁNDEZ, Roberto, FERNANDEZ, Carlos, BAPTISTA, Pilar
Metodología de la Investigación. México, McGraw-Hill, 2002.

- INGRAM, Elizabeth.
“Basic Concepts in Testing.” *The Edinburgh Course in Applied Linguistics*. Vol. 4. OUP.
London, 1977.

- Junta de Andalucía. “La evaluación educativa ” en *Relieve. Revista Electrónica de Investigación y Educación Educativa*, Núm. 1. España: Consejería de educación y Ciencia, diciembre, 1996.
<http://www.juntadeandalucia.es/educacion> (Consulta Septiembre 2005).

- KARMEL, Louis
Medición y Evaluación Escolar. México, Ed. Trillas, 1984.

- KIM, Hyun, SHIN Hye
“A Reading and Writing Placement Test: Design, Evaluation, and Analysis” en *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*”, Vol. 6, No. 2, 2006.

- LADO, Robert
Language Testing. The Construction and Use of Foreign Language Tests. New York, 1964.

- *Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación*. Madrid: Ministerio de Educación, Cultura y Deporte, 2002.
http://cvc.cervantes.es/obref/marco/cvc_mer.pdf (Consulta Septiembre 2005).

- McDONALD, Mary
Systematic Assessment of Learning Outcomes: Developing Multiple Choice Exams.
Boston, Jones and Bartlett, 2001.
http://books.google.com.mx/books?id=MEPZD5ALOTIC&printsec=titlepage&dq=mean+biserial&source=gbs_toc_s&cad=1#PPA7,M1 (Consulta Mayo 2008).

- MURRAY, Joel R.
Entrance Testing at a Private English Language Institute: Toward More Accurate Placements. University of British Columbia, 1999. (Tesis no publicada).

- MURRAY, Joel R.
Steps and Recommendations for More Accurate Placement Test Creation
U.S. Department of Education. Educational Resources Information Center (ERIC). ED 453 656 (2001).
http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/17/15/eb.pdf (Consulta marzo 2006).

- O’SULLIVAN, Barry, Cyrill Weyr, y Nick Saville
“Using observation checklists to validate speaking-test tasks” en *Language Testing*, Vol. 19, No. 1, 2002.

- PARK, Taejoon
“An Investigation of an ESL Placement Test of Writing Using Many-facet Rasch Measurement.” en *Working Papers in TESOL & Applied Linguistics*, Vol. 4, Núm 1, Teachers College, Columbia University, 2004.

- RUEDA, Mario, y Javier Nieto

La Evaluación de la Docencia Universitaria. México: CISE, UNAM, 1994.

- WALL, Diane, Caroline Clapham , Charles Alderson

Evaluating a placement test. *Language Testing*. Vol. 11, No. 3. SAGE Publications, 1994.

(Consulta Junio, 2008). <http://ltj.sagepub.com/cgi/content/abstract/11/3/321>

APÉNDICES

1. Hoja de información del aspirante.
2. Cuestionario de salida para aspirantes.
3. Cuestionario para profesores.
4. Cuestionario para alumnos.
5. Matriz Conducta / Contenido.
6. Tabla de especificaciones.
7. Análisis de reactivos con *ITEMAN 2.0*.

Apéndice 1



CENTRO DE IDIOMAS INFORMACIÓN DEL ASPIRANTE

LA PRESENTE INFORMACIÓN ES CONFIDENCIAL

El objetivo de obtener los siguientes datos es contar con información estadística sobre los aspirantes a los cursos de Inglés del Centro de Idiomas de la UIA

Le agradecemos el tiempo y los detalles que se sirva proporcionar.

Nombre: _____
Teléfono: _____ E-mail _____
Edad: _____

1. Elija la opción que corresponda:

Alumno UIA Candidato externo

2. Alumno UIA: Carrera: _____ Semestre: _____

Candidato externo: Ocupación: _____

3. Razones para realizar el presente proceso de clasificación (marque las que apliquen por orden de importancia)
1 (+ importante), 2 (regularmente importante), 3 (menos importante)

- a. Requisito idioma UIA
- b. Superación personal en el idioma
- c. Superación laboral

4. Trayectoria escolar (marque lo que corresponda):

- a. Primaria Escuela privada: Escuela Pública:
Nombre de la escuela: _____
- b. Secundaria Escuela privada: Escuela Pública:
Nombre de la escuela: _____
- c. Preparatoria Escuela privada: Escuela Pública:
Nombre de la escuela: _____

- d. Cursos en institutos de idiomas
Nombre: _____
Nombre: _____
Nombre: _____
- e. ¿Cuándo realizó sus últimos estudios formales en el idioma)? Hace _____ años

f. Con base en la tabla adjunta, indique cuál considera que es su nivel de idioma actualmente:
Nivel: _____

5. Exprese cualquier comentario que considere pertinente sobre su conocimiento del idioma inglés

Nota: Una vez que conozca su nivel definitivo le agradeceremos regresar a este salón

CUADRO DE AUTOEVALUACIÓN

1) Principiante	<p>Expresión escrita: Soy capaz de escribir postales cortas y sencillas: por ejemplo para enviar felicitaciones. Se rellenan formularios con datos personales; por ejemplo mi nombre, mi nacionalidad y mi dirección en el formulario del registro de un hotel.</p> <p>Expresión Oral: Utilizo expresiones y frases sencillas para describir el lugar donde vivo y las personas que conozco.</p> <p>Comprensión de lectura: Comprendo palabras y nombres conocidos y frases muy sencillas; por ejemplo las que hay en letreros, carteles y catálogos.</p> <p>Comprensión auditiva: Reconozco palabras y expresiones muy básicas que se usan habitualmente, relativas a mi mismo, a mi familia y a mi entorno inmediato cuando se habla despacio y con claridad.</p>
2) Elemental	<p>Expresión escrita: Soy capaz de escribir notas y mensajes breves y sencillos relativos a mis necesidades inmediatas. Puedo escribir cartas personales muy sencillas: por ejemplo, agradeciendo algo a alguien.</p> <p>Expresión oral: Utilizo una serie de expresiones y frases para describir con términos sencillos a mi familia y a otras personas, mis condiciones de vida, mi origen educativo y mi trabajo actual, o el último que tuve.</p> <p>Comprensión de lectura: Soy capaz de leer textos muy breves y sencillos. Sé encontrar información específica y predecible en escritos sencillos y cotidianos como anuncios publicitarios, prospectos menús y horarios, y comprendo cartas personales breves y sencillas.</p> <p>Comprensión auditiva: Comprendo frases y el vocabulario más habitual sobre temas de interés personal (información personal y familiar muy básica, compras, lugar de residencia, empleo). Soy capaz de captar la idea principal de avisos y mensajes breves, claros y sencillos.</p>
3) Pre-intermedio	<p>Expresión escrita: Soy capaz de escribir textos sencillos y bien enlazados sobre temas que me son conocidos o de interés personal. Puedo escribir cartas personales que describen experiencias e impresiones.</p> <p>Expresión oral: Sé enlazar frases de forma sencilla con el fin de describir experiencia y hechos, mis sueños, esperanzas y ambiciones. Puedo explicar y justificar brevemente mis opiniones y proyectos. Sé narrar una historia o relato, la trama de un libro o película, y puedo describir mis reacciones.</p> <p>Comprensión de lectura: Comprendo textos redactados en una lengua de uso habitual y cotidiano o relacionada con el trabajo. Comprendo la descripción de acontecimientos, sentimientos y deseos en cartas personales.</p> <p>Comprensión auditiva: Comprendo las ideas principales cuando el discurso es claro y normal y se tratan asuntos cotidianos que tienen lugar en el trabajo, en la escuela, durante el tiempo de ocio, etc. Comprendo la idea principal de muchos programas de radio o televisión que tratan temas actuales o asuntos de interés personal o profesional, cuando la articulación es relativamente lenta y clara.</p>
4) Intermedio	<p>Expresión escrita: Soy capaz de escribir textos claros y detallados sobre una amplia serie de temas relacionados con mis intereses. Puedo escribir redacciones o informes transmitiendo información o proponiendo motivos que apoyen o refuten un punto de vista concreto. Sé escribir cartas que destacan la importancia que le doy a determinados hechos y experiencias.</p> <p>Expresión oral: Presento descripciones claras y detalladas de una amplia serie de temas relacionados con mi especialidad. Sé explicar un punto de vista sobre un tema exponiendo las ventajas y los inconvenientes de varias opciones.</p> <p>Comprensión de lectura: Soy capaz de leer artículos e informes relativos a problemas contemporáneos en los que los autores adoptan posturas o puntos de vista concretos. Comprendo la prosa literaria y contemporánea.</p> <p>Comprensión auditiva: Comprendo discursos y conferencias extensos e incluso sigo líneas argumentales complejas siempre que el tema sea relativamente conocido. Comprendo casi todas las noticias de la televisión y los programas sobre temas actuales. Comprendo la mayoría de las películas en las que se habla en un nivel de lengua estándar.</p>
5) Intermedio-Alto	<p>Expresión escrita: Soy capaz de expresarme en textos claros y bien estructurados, exponiendo puntos de vista con cierta extensión. Puedo escribir sobre temas complejos en cartas, redacciones o informes, resaltando lo que considero que son aspectos importantes. Seleccione el estilo apropiado para los lectores a los que van dirigidos mis escritos.</p> <p>Expresión Oral: Presento descripciones claras y detalladas sobre temas complejos que incluyen otros temas, desarrollando ideas concretas y terminando con una conclusión apropiada.</p> <p>Comprensión de lectura: Comprendo textos largos y complejos de carácter literario o basados en hechos, apreciando distinciones de estilo. Comprendo artículos especializados e instrucciones técnicas largas, aunque no se relacionen con mi especialidad.</p> <p>Comprensión auditiva: Comprendo discursos extensos incluso cuando no están estructurados con claridad y cuando las relaciones están sólo implícitas y no se señalan explícitamente. Comprendo sin mucho esfuerzo los programas de televisión y las películas.</p>
6) Avanzado	<p>Expresión escrita: Soy capaz de escribir textos claros y fluidos en un estilo apropiado. Puedo escribir cartas, informes o artículos complejos que presenten argumentos con una estructura lógica y eficaz. Escribo resúmenes y reseñas de obras profesionales o literarias.</p> <p>Expresión Oral: Presento descripciones o argumentos de forma clara y fluida, y con un estilo que es adecuado al contexto.</p> <p>Comprensión de lectura: Soy capaz de leer con facilidad prácticamente todas las formas de lengua escrita, incluyendo textos abstractos estructural o lingüísticamente complejos como, por ejemplo, manuales, artículos especializados y obras literarias.</p> <p>Comprensión auditiva: No tengo ninguna dificultad para comprender cualquier tipo de lengua hablada, tanto en conversaciones en vivo como en discursos retransmitidos, aunque se produzcan a una velocidad de hablante nativo, siempre que tenga tiempo para familiarizarme con el acento.</p>

Apéndice 3



UNIVERSIDAD
IBEROAMERICANA

CENTRO DE IDIOMAS CUESTIONARIO PARA PROFESORES

El presente cuestionario tiene como objetivo conocer sus impresiones con respecto al nivel de idioma de los estudiantes que participan en su curso durante el presente periodo.

Le agradecemos el tiempo y los detalles que se sirva proporcionar.

Nombre del profesor: _____

Nombre del curso: _____

Fecha: _____

- a. Número de alumnos inscritos inicialmente: _____
- b. Número de alumnos que a la fecha asisten a clase de manera regular: _____
- c. En el presente periodo ¿detectó usted algún alumno cuyo nivel no correspondiera al curso impartido?
Sí No En este caso, agradecemos su tiempo
- d. Si la respuesta es positiva indique el número de alumnos en dicha circunstancia: _____
- e. A continuación le agradeceremos proporcionar la siguiente información sobre dichos alumnos:

Nombre del alumno: _____
* Origen del alumno: Curso anterior <input type="checkbox"/> Proceso clasificación <input type="checkbox"/> Lo ignoro <input type="checkbox"/>
* Indique la problemática personal del alumno: Si su nivel era más bajo o más alto, y en cuáles de las habilidades lingüísticas: <u>producción</u> (oral o escrita), <u>comprensión</u> (oral o escrita): _____ _____
* Indique la problemática que le representó a nivel grupo la participación de este alumno en su clase: _____ _____ _____ _____
* Indique qué sucedió con el alumno: Cambió de grupo <input type="checkbox"/> Indique a qué curso se le cambió _____ Continúa asistiendo <input type="checkbox"/> De detalles sobre su desempeño: _____ _____ _____ No regresó <input type="checkbox"/> Si conoce alguna razón específica, indíquela: _____ _____ _____
* Por favor agregue cualquier comentario adicional que considere importante sobre el alumno: _____ _____ _____ _____

Apéndice 4



CENTRO DE IDIOMAS CUESTIONARIO PARA ESTUDIANTES INSCRITOS

El presente cuestionario tiene como objetivo conocer tus impresiones con respecto al proceso de clasificación del nivel de inglés que se utiliza en el Centro de Idiomas de la UIA.

Te agradecemos el tiempo para contestar las siguientes preguntas, así como los comentarios que nos puedas proporcionar.

Nombre del alumno: _____

Nombre del profesor: _____

Nombre del curso: _____

Fecha: _____

- a. ¿Cuántos cursos has realizado en el Centro de idiomas de la UIA? _____
- b. ¿Cuándo realizaste el proceso de clasificación (menciona año y periodo: primavera, verano, otoño)? _____
- c. ¿Consideras que cuando realizaste el proceso de clasificación se te ubicó en el nivel adecuado?:
Sí pasa a la pregunta "e" No
- d. Si la respuesta es negativa, explica por qué: _____

- e. Si hubieras repetido el proceso de clasificación al siguiente día de que lo hiciste (sin estudiar para llevarlo a cabo) ¿crees que el resultado hubiera determinado el mismo nivel de idioma?

Sí pasa a la pregunta "g" No

- f. Si la respuesta es negativa explica por qué (¿Existió algún factor que no te permitiera dar lo mejor de ti mismo al realizar el proceso (ejemplo: enfermedad, ruido, presión, etc.)? _____

- g. Una vez inscrito en el nivel determinado ¿consideras que tu nivel para participar en el curso fue el adecuado?

Sí pasa a la siguiente página No

- h. Si la respuesta es negativa, explica por qué [ejemplo: ¿tu nivel era más alto o más bajo que el del resto del grupo? Si así fue ¿en cuál o cuáles habilidades lingüísticas: producción (oral o escrita), comprensión (oral o escrita)]:

- i. ¿Fuiste cambiado de nivel?

Sí No pasa a la pregunta "l"

- j. En caso positivo ¿a cuál curso? _____

- k. ¿Cómo te sentiste con el cambio y por qué? _____

pasa a la pregunta "m"

- l. ¿Por qué? (ej: no lo comentaste con el profesor o en la coordinación; el profesor no se percató; no te pareció importante; estabas a gusto con el grupo; preferiste repasar; otra situación) _____

- m. Si el día de hoy realizaras nuevamente el proceso de clasificación ¿crees que te ubicarían en el nivel inmediato superior como en teoría corresponde, o que podrías dejar de realizar por lo menos el siguiente nivel? _____

- n. En caso de que pudieras saltar un nivel de estudio ¿a qué lo atribuirías? (ej: a que en un principio estuviste mal clasificado; a que además de tu actual curso has estudiado por tu cuenta; a que al realizar el curso recordaste cosas que ya habías estudiado; a alguna otra causa -¿cuál?)

- o. Por favor expresa cualquier otro comentario o sugerencia que tengas sobre el proceso de clasificación de nivel del idioma

Apéndice 5

1/6

Principiante

MATRIZ CONDUCTA-CONTENIDO

CONDUCTA CONTENIDO	1	2	3 P 06	3 Ideal	4	5	6	%	No. de reactivos
Principiante 1									
Auxiliar 'be' <i>am, are, is</i> [+, -, ? (-wh & yes/no)]			2	2				13	2
Possesive adjectives (<i>my, your, his, her, our, their</i>)			2	1				6.7	2
Object Pronouns (<i>me, you, him, her, it, us, them</i>)				1				6.7	
Posesivo ('s)			1	1				6.7	1
Have & has			1	1				6.7	1
Presente simple [+, -, ? (-wh & yes/no)]			4	4				26.7	4
artículos indefinidos <i>a/an</i>			1	1				6.7	1
Adverbios de frecuencia (<i>never, sometimes, usually</i>)			1	1				6.7	1
<i>this/that</i>				1				6.7	
<i>There is/there are</i>				1				6.7	
<i>Any (-, ?)</i>			1	1				6.7	1
%				100				100	
No. de reactivos			13	13					15
Principiante 2									
Auxiliar 'be' <i>was, were</i> [+, -, ? (-wh & yes/no)]			1	1				6.7	1
Pasado simple [+, -, ? (-wh & yes/no)]			7	5				47	7
Auxiliar <i>can</i> [+, -, ? (-wh & yes/no)]			1	1				6.7	1
Like + <i>ing</i>			1	1				6.7	1
Would like & want			1	1				6.7	1
Presente continuo [+, -, ? (-wh & yes/no)]			2	2				13	2
Presente simple vs. Continuo				2					
Presente continuo (para futuro)			2	2				13	2
%			100	100				99.8	
No. de reactivos			15	15					15

Elemental

MATRIZ CONDUCTA-CONTENIDO

<i>CONDUCTA</i>	1	2	3 <i>P 06</i>	3 <i>Ideal</i>	4	5	6	%	<i>No. de reactivos</i>
<i>CONTENIDO</i>									

Elemental 3									
Auxiliar 'be' am, are, is [+, -, ? (-wh & yes/no)]			2	?				13	2
Possesive adjectives (my, your, his, her, our, their)				?					
Presente simple [+, -, ? (-wh & yes/no)]			5	3				33.3	5
There is/there are				1					
Adverbios de frecuencia (never, sometimes, often, usually, always)			1	1				6.7	1
Preposiciones de lugar (in, on, at, in front of, under, next to)			3	1				20	3
How many, ..., some (+), any (-, ?)				1					
this, that, these, those			1	1				6.7	1
can, can't (habilidad en presente)				1					
Pasado 'be' was, were [+, -, ? (-wh & yes/no)]			1	1				6.7	1
could, couldn't (habilidades en el pasado)				1					
Pasado simple [+, -, ? (-wh & yes/no)]			2	3				13	2
Ago & Time expressions (in, on, at)				1					
%			100	100				99.4	
No. de reactivos			15	15					15
Elemental 4									
Sustantivos contables e incontables				1					
Would you like..? vs. Do you like..?			2	1				13	2
Artículo a, & some, any									
much y many									
Comparativos y superlativos			3	2				20	3
Preposiciones de lugar (in, on, at, in front of, under, next to)				?					
Presente simple y continuo			1	1				6.7	1
Whose' para posesivos				1					
Pronombres posesivos (mine, yours, his, hers, ours, theirs)				1					
Futuro idiomático going to...			1	2				6.7	1
Infinitivo para expresar propósito				1					
Adjetivos y adverbios (quick, quickly, good, well, etc)			2	1				13	2
Presente perfecto			1	1				6.7	1
ever & never			1	1				6.7	1
yet & just			2	1				13	2
Presente perfecto y pasado simple			2	1				13	2
%			100	100				98.8	
No. de reactivos			15	15					15

Ap5 - 3/6

Pre-Intermedio

MATRIZ CONDUCTA-CONTENIDO

CONDUCTA	1	2	3	3	4	5	6	%	No. de reactivos
CONTENIDO			P 06	Ideal					

Pre-Intermedio 5								
preguntas simples y abiertas								
Presente simple y presente cont.			2				13	2
Have/has & have/has got								
Pasado simple y pasado cont.			2				13	2
Preposiciones de tiempo (in, at, on)								
much, many, some, any,			2				13	2
someone, anyone, somewhere, etc								
a few, a little a lot/ lots of								
Articulos a, an, the			1				6.7	1
Patrones verbales			1				6.7	1
Will para decisiones espontáneas								
Going to... para planes y predicciones (evidentes)			1				6.7	1
Linking words (but, however, so, because)			1				6.7	1
What...like? para descripciones			1				6.7	1
How is...? para salud o edo. de animo								
Comparativos y superlativos			1				6.7	1
Presente perfecto y Pasado simple			3				20	3
%			100				99.2	
No. de reactivos			15					15
Pre-Intermedio 6								
have to vs. have got to								
Modal verbs (can, could, might, must, should, will, would)			2				13	2
Time clauses (while, when, as soon as, after, before, until)			1				6.7	1
1er condicional			1				6.7	1
Phrasal verbs			1				6.7	1
Patrones verbales			1				6.7	1
Used to y pasado simple			1				6.7	1
Usos del infinitivo								
Voz pasiva: presente, pasado, presente perfecto, futuro			2				13	2
2o. Condicional			1				6.7	1
Presente perfecto continuo			2				13	2
Pasado perfecto			1				6.7	1
Reported statements			2				13	2
%			100				98.9	
No. de reactivos			15					15

Ap5 - 4/6

Intermedio

MATRIZ CONDUCTA-CONTENIDO

CONDUCTA CONTENIDO	1	2	3 <i>P 06</i>	3 <i>Ideal</i>	4	5	6	%	No. de reactivos
Intermedio 7									
Tiempos presentes (simple, continuo, perfecto)			2						2
Voz pasiva, presente, pasado, pres.pfcto (simple y continuo)			3						3
Pasado simple y continuo			1						1
Used to			1						1
Pasado simple y perfecto			2						2
Modal verbs (be allowed to, have to, should, must)			1						1
Futuro: <i>going to ... Will, Presente continuo</i>			2						2
like, what...like?, look like									
Verbos estáticos			1						1
Relative Clauses			2						2
Patrones verbales									
%			100						
No. de reactivos			15						15
Intermedio 8									
Presente perfecto y pasado simple (simple y continuo)			3						3
Voz pasiva: presente perfecto			1						1
1o. y 2o. condicional			3						3
Time clauses (while, when, as soon as, after, before, until)									
Modal verbs en pasado (must've, could've, might've, cant have)			1						1
Presente perfecto simple vs. cont.									
Indirect speech			1						1
Question tags									
Reported speech (afirmaciones, preguntas, peticiones/órdenes)			1						1
Agreeing/disagreeing (<i>so do I, neither do I, etc</i>)			2						2
<i>Would</i> para preferencias			1						1
Modifying adverbs (very, absolutely, etc)			1						1
%									
No. de reactivos			14						15

Ap-5 - 5/6

Intermedio-Alto

MATRIZ CONDUCTA-CONTENIDO

CONDUCTA CONTENIDO	1	2	3 <i>P 06</i>	3 <i>Ideal</i>	4	5	6	%	No. de reactivos
Intermedio alto 9									
Voz activa y pasiva (todos los tiempos)									
Presente perfecto y continuo			3						
Pasado simple, continuo y perfecto									
Time clauses (while, when, as soon as, after, before, until)									
Quantifiers (<i>some/any, much/many, each/every, more/most, a little/little, a few/few, both, fewer/less, several, all/no, enough, none, plenty of, etc.</i>)			2						
Futuro: <i>will, going to, presente continuo, simple will y shall</i>									
Futuro continuo									
Futuro perfecto									
Cláusulas relativas									
Participios (presente y pasado)									
%									
No. de reactivos									15
Intermedio alto 10									
Patrones verbales									
Patrones verbales e infinitivo									
Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos (<i>willingnes-refusal</i>)									
Preguntas con <i>what</i> y <i>which</i> , para descripción e indirectas									
Uso de negativos									
Hábitos con Presente simple, continuo, <i>will, would, used to, be/get used to</i>									
1o. y 2o. condicional									
3er condicional									
<i>I wish, If only</i> para hipótesis									
Frases sustantivas (<i>noun phrases</i>)									
Artículos <i>a, an, the, zero article</i>									
Énfasis <i>Word order</i> , estructuras enfáticas									
%									
No. de reactivos									15

Avanzado

Ap5 - 6/6

MATRIZ CONDUCTA-CONTENIDO

CONDUCTA CONTENIDO	1	2	3 P 06	3 Ideal	4	5	6	%	No. de reactivos
Avanzado 11									
Expresiones plurales con adjetivos compuestos									
Auxiliares para evitar repetición									
Artículos <i>a, an, the</i>									
Pasados: simple, continuo, perfecto, futuro en el pasado									
as vs. like									
Futuro: <i>will, going to, presente continuo, presente simple (p/itinerarios); futuro perfecto</i>									
Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos									
Orden de adjetivos									
%									
No. de reactivos									15
Avanzado 12									
Patrones verbales en Reported Speech									
Condicionales (3 tipos y mixed)									
Cláusulas de tiempo en pasado									
As; 'while', 'when', 'after', 'before'									
Cláusulas de tiempo en futuro									
Cláusulas relativas (defining & non-defining)									
Participios (como adjetivos, como cláusulas relativas reducidas, en cláusulas adverbiales, después de ciertos verbos.									
Inversión para expresar énfasis									
Hipótesis (presente y futuro: 'I wish' 'if only', hechos vs. suposiciones: 'it's time'; 'would rather')									
Hipótesis (pasada)									
Would like'									
Presente subjuntivo (después de verbos, adj., sustantivos; después de patrones verbales). Alternativas de subjuntivo: should									
Pasado subjuntivo									
%									
No. de reactivos									15
%		26.92		45.15	7.69	9.61	9.61	100	
No. de reactivos		14		24	4	5	5		52

Apéndice 6

1/6

Principiante

TABLA DE ESPECIFICACIONES

No. de Pregunta	Contenido	Nivel Taxonómico	Tipo de Reactivo	Valor del reactivo	OBSERVACIONES
Principiante 1					
1	Auxiliar 'be' <i>am, are, is (+, -, ?)</i>	3	Opción múltiple	1	
2	Auxiliar 'be' <i>am, are, is (+, -, ?)</i>	3	Opción múltiple	1	
3	artículos indefinidos <i>a/an</i>	3	Opción múltiple	1	
4	<i>Any (-, ?)</i>	3	Opción múltiple	1	
5	Have & has	3	Opción múltiple	1	
6	Possesive adjectives (<i>my, your, his, her, our, their</i>)	3	Opción múltiple	1	
7		3	Opción múltiple	1	
8	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
9	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
10	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
11	Posesivo ('s)	3	Opción múltiple	1	
12	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
13		3	Opción múltiple	1	
14	Adverbios de frecuencia (<i>never, sometimes, usually</i>)	3	Opción múltiple	1	
15	Possesive adjectives (<i>my, your, his, her, our, their</i>)	3	Opción múltiple	1	
Principiante 1					
16	Like + <i>ing</i>	3	Opción múltiple	1	
17	Auxiliar <i>can (+, -, ?)</i>	3	Opción múltiple	1	
18	Auxiliar 'be' <i>was, were (+, -, ?)</i>	3	Opción múltiple	1	
19	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
20	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
21	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
22	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
23	Presente continuo [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
24	Presente continuo [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
25	Presente continuo (para futuro)	3	Opción múltiple	1	
26	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
27	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
28	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
29	Presente continuo (para futuro)	3	Opción múltiple	1	
30	Would like & want	3	Opción múltiple	1	

Elemental

TABLA DE ESPECIFICACIONES

No. de Pregunta	Contenido	Nivel Taxonómico	Tipo de Reactivo	Valor del reactivo	OBSERVACIONES
Elemental 3					
1	Preposiciones de lugar (<i>in, on, at, in front of, under, next to</i>)	3	Opción múltiple	1	
2	Auxiliar 'be' <i>am, are, is</i> [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
3	Preposiciones de lugar (<i>in, on, at, in front of, under, next to</i>)	3	Opción múltiple	1	
4	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
5	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
6	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
7	Auxiliar 'be' <i>am, are, is</i> [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
8	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
9	Adverbios de frecuencia (<i>never, sometimes, often, usually, always</i>)	3	Opción múltiple	1	
10	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
11	Preposiciones de lugar (<i>in, on, at, in front of, under, next to</i>)	3	Opción múltiple	1	
12	Presente simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
13	<i>this, that, these, those</i>	3	Opción múltiple	1	
14	Pasado 'be' <i>was, were</i> [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
15	Pasado simple [+, -, ? (-wh & yes/no)]	3	Opción múltiple	1	
Elemental 4					
16	Futuro idiomático <i>going to...</i>	3	Opción múltiple	1	
17	Presente perfecto y pasado simple	3	Opción múltiple	1	
18	Adjetivos y adverbios (<i>quick, quickly, good, well, etc</i>)	3	Opción múltiple	1	
19	Presente perfecto y pasado simple	3	Opción múltiple	1	
20	<i>Would you like..? vs. Do you like..?</i>	3	Opción múltiple	1	
21	<i>Would you like..? vs. Do you like..?</i>	3	Opción múltiple	1	
22	Comparativos y superlativos	3	Opción múltiple	1	
23	Comparativos y superlativos	3	Opción múltiple	1	
24	Comparativos y superlativos	3	Opción múltiple	1	
25	Presente simple y continuo	3	Opción múltiple	1	
26	<i>yet & just</i>	3	Opción múltiple	1	
27	Presente perfecto	3	Opción múltiple	1	
28	<i>yet & just</i>	3	Opción múltiple	1	
29	Adjetivos y adverbios (<i>quick, quickly, good, well, etc</i>)	3	Opción múltiple	1	
30	<i>ever & never</i>	3	Opción múltiple	1	

Pre-Intermedio

TABLA DE ESPECIFICACIONES

No. de Pregunta	Contenido	Nivel Taxonómico	Tipo de Reactivo	Valor del reactivo	OBSERVACIONES
Pre-Intermedio 5					
1	Presente simple y presente cont.	3	Opción múltiple	1	
2	<i>much, many, some, any,</i>	3	Opción múltiple	1	
3	Presente simple y presente cont.	3	Opción múltiple	1	
4	Pasado simple y pasado cont.	3	Opción múltiple	1	
5	<i>What...like?</i> para descripciones	3	Opción múltiple	1	
6	Patrones verbales	3	Opción múltiple	1	
7	<i>Will</i> para decisiones espontáneas	3	Opción múltiple	1	
8		3	Opción múltiple	1	
9	<i>much, many, some, any,</i>	3	Opción múltiple	1	
10	Presente perfecto y Pasado simple	3	Opción múltiple	1	
11	Presente perfecto y Pasado simple	3	Opción múltiple	1	
12	Presente perfecto y Pasado simple	3	Opción múltiple	1	
13	<i>Going to...</i> para planes y predicciones (evidentes)	3	Opción múltiple	1	
14	<i>Comparativos y superlativos</i>	3	Opción múltiple	1	
15	Pasado simple y pasado cont.	3	Opción múltiple	1	
Pre-Intermedio 6					
16	<i>Patrones verbales</i>	3	Opción múltiple	1	
17	Phrasal verbs	3	Opción múltiple	1	
18	<i>1er condicional</i>	3	Opción múltiple	1	
19	Presente perfecto continuo	3	Opción múltiple	1	
20	2o. Condicional	3	Opción múltiple	1	
21	Time clauses (<i>while, when, as soon as, after, before, until</i>)	3	Opción múltiple	1	
22	<i>Used to</i> y pasado simple	3	Opción múltiple	1	
23	Modal verbs (<i>can, could, might, must, should, will, would</i>)	3	Opción múltiple	1	
24	Reported statements	3	Opción múltiple	1	
25	Voz pasiva: presente, pasado, presente perfecto, futuro	3	Opción múltiple	1	
26	Modal verbs (<i>can, could, might, must, should, will, would</i>)	3	Opción múltiple	1	
27	Reported statements	3	Opción múltiple	1	
28	Presente perfecto continuo	3	Opción múltiple	1	
29	Voz pasiva: presente, pasado, presente perfecto, futuro	3	Opción múltiple	1	
30	Pasado perfecto	3	Opción múltiple	1	

Intermedio

TABLA DE ESPECIFICACIONES

No. de Pregunta	Contenido	Nivel Taxonómico	Tipo de Reactivo	Valor del reactivo	OBSERVACIONES
Intermedio 7					
1	Pasado simple y perfecto	3	Opción múltiple	1	
2	Tiempos presentes (simple, continuo, perfecto)	3	Opción múltiple	1	
3	Voz pasiva, presente, pasado, pres.pfcto (simple y continuo)	3	Opción múltiple	1	
4	Tiempos presentes (simple, continuo, perfecto)	3	Opción múltiple	1	
5	Verbos estáticos	3	Opción múltiple	1	
6	Voz pasiva, presente, pasado, pres.pfcto (simple y continuo)	3	Opción múltiple	1	
7	Pasado simple y continuo	3	Opción múltiple	1	
8	Voz pasiva, presente, pasado, pres.pfcto (simple y continuo)	3	Opción múltiple	1	
9	Modal verbs (be allowed to, have to, should, must)	3	Opción múltiple	1	
10	Used to	3	Opción múltiple	1	
11	Relative Clauses	3	Opción múltiple	1	
12	Relative Clauses	3	Opción múltiple	1	
13	Pasado simple y perfecto	3	Opción múltiple	1	
14	Futuro: <i>going to ... Will, Presente continuo</i>	3	Opción múltiple	1	
15	Futuro: <i>going to ... Will, Presente continuo</i>	3	Opción múltiple	1	
Intermedio 8					
16	Modifying adverbs (very, absolutely, etc)	3	Opción múltiple	1	
17	1o. y 2o. condicional	3	Opción múltiple	1	
18	1o. y 2o. condicional	3	Opción múltiple	1	
19	Indirect speech	3	Opción múltiple	1	
20	1o. y 2o. condicional	3	Opción múltiple	1	
21	Presente perfecto y pasado simple (simple y continuo)	3	Opción múltiple	1	
22	Presente perfecto y pasado simple (simple y continuo)	3	Opción múltiple	1	
23	Agreeing/disagreeing (<i>so do I, neither do I, etc</i>)	3	Opción múltiple	1	
24	Presente perfecto y pasado simple (simple y continuo)	3	Opción múltiple	1	
25	Agreeing/disagreeing (<i>so do I, neither do I, etc</i>)	3	Opción múltiple	1	
26	1o. y 2o. condicional	3	Opción múltiple	1	
27	<i>Would</i> para preferencias	3	Opción múltiple	1	
28	Indirect speech	3	Opción múltiple	1	
29	Modal verbs en pasado (must've, could've, might've, cant have)	3	Opción múltiple	1	
30		3	Opción múltiple	1	

Intermedio-Alto

TABLA DE ESPECIFICACIONES

No. de Pregunta	Contenido	Nivel Taxonómico	Tipo de Reactivo	Valor del reactivo	OBSERVACIONES
Intermedio alto 9					
1		3	Opción múltiple	1	Patrones verb e infinitivo INT ALTO 10
2	Presente perfecto y continuo	3	Opción múltiple	1	
3	Presente perfecto y continuo	3	Opción múltiple	1	
4	Presente perfecto y continuo	3	Opción múltiple	1	
5	<i>Quantifiers (both)</i>	3	Opción múltiple	1	
6	<i>Quantifiers (plenty of)</i>	3	Opción múltiple	1	
7	Pasado simple, continuo y perfecto	3	Opción múltiple	1	
8		3	Opción múltiple	1	willinness-refusal INT ALTO 10
9	Cláusulas relativas	3	Opción múltiple	1	
10	Futuro continuo	3	Opción múltiple	1	
11	Participios (presente y pasado)	3	Opción múltiple	1	
12		3	Opción múltiple	1	Condicionales INT ALTO 10
13	<i>Futuro: will, going to, presente continuo, simple</i>	3	Opción múltiple	1	
14	<i>Quantifiers (a little/little, a few/few)</i>	3	Opción múltiple	1	
15	Pasado simple, continuo y perfecto	3	Opción múltiple	1	
Intermedio alto 10					
16	Hábitos con <i>used to</i>	3	Opción múltiple	1	
17	<i>Modal auxiliaries</i> , probabilidad pasadal)	3	Opción múltiple	1	
18	<i>I wish, If only</i> para hipótesis	3	Opción múltiple	1	
19	<i>I wish, If only</i> para hipótesis	3	Opción múltiple	1	
20	<i>Modal auxiliaries</i> . Otros usos (willingness-refusal)	3	Opción múltiple	1	
21	<i>Modal auxiliaries</i>	3	Opción múltiple	1	
22	Patrones verbales	3	Opción múltiple	1	
23	Patrones verbales	3	Opción múltiple	1	
24	Hábitos con <i>be/get used to</i>	3	Opción múltiple	1	
25	<i>I wish, If only</i> para hipótesis	3	Opción múltiple	1	
26		3	Opción múltiple	1	Participio presente INT ALTO 9
27	Patrones verbales	3	Opción múltiple	1	
28		3	Opción múltiple	1	VOZ PASIVA INT ALTO 9
29	2o. condicional	3	Opción múltiple	1	
30	2o. condicional	3	Opción múltiple	1	

Avanzado

TABLA DE ESPECIFICACIONES

No. de Pregunta	Contenido	Nivel Taxonómico	Tipo de Reactivo	Valor del reactivo	OBSERVACIONES
Avanzado 11					
1	Expresiones plurales con adjetivos compuestos	3	Opción múltiple	1	
2	Pasados: simple, continuo, perfecto, futuro en el pasado	3	Opción múltiple	1	
3	Pasados: simple, continuo, perfecto, futuro en el pasado	3	Opción múltiple	1	
4	Pasados: simple, continuo, perfecto, futuro en el pasado	3	Opción múltiple	1	
5	as vs. like	3	Opción múltiple	1	
6	<i>Futuro: presente simple (p/itinerarios);</i>	3	Opción múltiple	1	
7	<i>Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos</i>	3	Opción múltiple	1	
8	<i>Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos</i>	3	Opción múltiple	1	
9	<i>Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos</i>	3	Opción múltiple	1	
10		3	Opción múltiple	1	No pertenece
11	<i>Futuro: perfecto</i>	3	Opción múltiple	1	
12	Orden de adjetivos	3	Opción múltiple	1	
13	Orden de adjetivos	3	Opción múltiple	1	
14	<i>Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos</i>	3	Opción múltiple	1	
15	<i>Modal auxiliaries, probabilidad pasada, presente y futura. Otros usos</i>	3	Opción múltiple	1	
Avanzado 12					
16	<i>Presente subjuntivo (después de verbos, adj., sustantivos y patrones verbales). Alternativas de subjuntivo: should</i>	3	Opción múltiple	1	
17	<i>Condicionales (3 tipos y mixed)</i>	3	Opción múltiple	1	
18	<i>Cláusulas de tiempo en futuro</i>	3	Opción múltiple	1	
19	<i>Cláusulas de tiempo en pasado</i>	3	Opción múltiple	1	
20	<i>Cláusulas relativas (defining & non-defining)</i>	3	Opción múltiple	1	
21	<i>Inversión para expresar énfasis</i>	3	Opción múltiple	1	
22	<i>Inversión para expresar énfasis</i>	3	Opción múltiple	1	
23	<i>Inversión para expresar énfasis</i>	3	Opción múltiple	1	
24	<i>Hipótesis (presente y futuro: 'I wish')</i>	3	Opción múltiple	1	
25	<i>Hipótesis (presente y futuro: 'it's time')</i>	3	Opción múltiple	1	
26	<i>Presente subjuntivo (después de sustantivos)</i>	3	Opción múltiple	1	
27	<i>Presente subjuntivo (después de adj.)</i>	3	Opción múltiple	1	
28	<i>Hipótesis ('would rather')</i>	3	Opción múltiple	1	
29	<i>Condicionales (3 tipos y mixed)</i>	3	Opción múltiple	1	
30	<i>Patrones verbales en Reported Speech</i>	3	Opción múltiple	1	

Apéndice 7

Ap7 - 1/27

Análisis de Reactivos con ITEMAN 2.01

MicroCAT (tm) Testing System
Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01
Item analysis for data from file lenguas.dat Page 1

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
1	0-1	0.688	0.246	0.188	1	0.075	-0.216	-0.116	
					2	0.131	-0.174	-0.109	
					3	0.688	0.246	0.188	*
					4	0.084	-0.040	-0.022	
					Other	0.022	-0.257	-0.091	
2	0-2	0.875	0.360	0.225	1	0.875	0.360	0.225	*
					2	0.047	0.014	0.006	
					3	0.057	-0.477	-0.236	
					4	0.020	-0.367	-0.126	
					Other	0.002	-0.575	-0.078	
3	0-3	0.790	0.393	0.278	1	0.790	0.393	0.278	*
					2	0.145	-0.163	-0.105	
					3	0.034	-0.607	-0.253	
					4	0.025	-0.444	-0.166	
					Other	0.005	-0.270	-0.057	
4	0-4	0.703	0.573	0.434	1	0.075	-0.542	-0.292	
					2	0.703	0.573	0.434	*
					3	0.066	-0.430	-0.222	
					4	0.152	-0.255	-0.167	
					Other	0.004	-0.559	-0.100	
5	0-5	0.591	0.539	0.426	1	0.039	-0.427	-0.187	
					2	0.591	0.539	0.426	*
					3	0.106	-0.561	-0.333	
					4	0.258	-0.190	-0.140	
					Other	0.005	-0.600	-0.127	
6	0-6	0.694	0.659	0.502	1	0.694	0.659	0.502	*
					2	0.100	-0.290	-0.170	
					3	0.054	-0.591	-0.286	
					4	0.122	-0.464	-0.287	
					Other	0.030	-0.318	-0.128	
7	0-7	0.756	0.568	0.415	1	0.109	-0.542	-0.325	
					2	0.054	-0.378	-0.183	
					3	0.756	0.568	0.415	*
					4	0.072	-0.211	-0.112	
					Other	0.009	-0.257	-0.066	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 2

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				Key
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	
8	0-8	0.767	0.709	0.513	1	0.068	-0.454	-0.237	
					2	0.767	0.709	0.513	*
					3	0.043	-0.193	-0.087	
					4	0.115	-0.672	-0.408	
					Other	0.007	-0.466	-0.110	
9	0-9	0.667	0.492	0.379	1	0.201	-0.508	-0.356	
					2	0.667	0.492	0.379	*
					3	0.116	-0.077	-0.047	
					4	0.004	-0.250	-0.045	
					Other	0.013	-0.563	-0.164	
10	0-10	0.774	0.557	0.401	1	0.774	0.557	0.401	*
					2	0.045	-0.322	-0.147	
					3	0.091	-0.376	-0.215	
					4	0.079	-0.442	-0.241	
					Other	0.011	-0.362	-0.100	
11	0-11	0.783	0.488	0.347	1	0.095	-0.417	-0.240	
					2	0.783	0.488	0.347	*
					3	0.027	-0.361	-0.139	
					4	0.084	-0.239	-0.133	
					Other	0.011	-0.474	-0.131	
12	0-12	0.803	0.527	0.368	1	0.081	-0.520	-0.286	
					2	0.052	-0.355	-0.170	
					3	0.803	0.527	0.368	*
					4	0.056	-0.143	-0.070	
					Other	0.009	-0.604	-0.155	
13	0-13	0.466	0.215	0.171	1	0.466	0.215	0.171	*
					2	0.052	-0.125	-0.060	
					3	0.068	-0.231	-0.121	
					4	0.385	-0.039	-0.030	
					Other	0.029	-0.410	-0.161	
14	0-14	0.690	0.637	0.486	1	0.134	-0.398	-0.253	
					2	0.116	-0.324	-0.198	
					3	0.690	0.637	0.486	*
					4	0.052	-0.585	-0.280	
					Other	0.007	-0.642	-0.151	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 3

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
15	0-15	0.819	0.737	0.504	1	0.039	-0.464	-0.203	
					2	0.819	0.737	0.504	*
					3	0.030	-0.293	-0.117	
					4	0.104	-0.675	-0.399	
					Other	0.007	-0.638	-0.150	
16	0-16	0.367	0.559	0.437	1	0.367	0.559	0.437	*
					2	0.398	-0.383	-0.302	
					3	0.047	-0.237	-0.110	
					4	0.163	-0.034	-0.023	
					Other	0.025	-0.535	-0.200	
17	0-17	0.697	0.310	0.235	1	0.697	0.310	0.235	*
					2	0.194	-0.008	-0.006	
					3	0.082	-0.436	-0.241	
					4	0.011	-0.458	-0.126	
					Other	0.016	-0.661	-0.211	
18	0-18	0.679	0.338	0.259	1	0.145	-0.110	-0.071	
					2	0.091	-0.204	-0.116	
					3	0.679	0.338	0.259	*
					4	0.061	-0.182	-0.092	
					Other	0.023	-0.731	-0.267	
19	0-19	0.550	0.480	0.382	1	0.231	-0.063	-0.045	
					2	0.550	0.480	0.382	*
					3	0.142	-0.356	-0.229	
					4	0.048	-0.461	-0.216	
					Other	0.029	-0.679	-0.267	
20	0-20	0.647	0.650	0.505	1	0.082	-0.522	-0.289	
					2	0.190	-0.195	-0.135	
					3	0.647	0.650	0.505	*
					4	0.052	-0.567	-0.272	
					Other	0.029	-0.745	-0.292	
21	0-21	0.523	0.369	0.294	1	0.387	-0.077	-0.061	
					2	0.023	-0.372	-0.136	
					3	0.038	-0.619	-0.267	
					4	0.523	0.369	0.294	*
					Other	0.029	-0.704	-0.277	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 4

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
22	0-22	0.846	0.712	0.468	1	0.846	0.712	0.468	*
					2	0.034	-0.181	-0.075	
					3	0.070	-0.573	-0.302	
					4	0.022	-0.557	-0.198	
					Other	0.029	-0.759	-0.298	
23	0-23	0.516	0.506	0.404	1	0.075	-0.551	-0.297	
					2	0.516	0.506	0.404	*
					3	0.217	-0.103	-0.074	
					4	0.138	-0.114	-0.073	
					Other	0.054	-0.623	-0.302	
24	0-24	0.783	0.629	0.448	1	0.041	-0.441	-0.196	
					2	0.070	-0.456	-0.240	
					3	0.075	-0.224	-0.121	
					4	0.783	0.629	0.448	*
					Other	0.030	-0.767	-0.307	
25	0-25	0.588	0.543	0.430	1	0.161	-0.538	-0.358	
					2	0.588	0.543	0.430	*
					3	0.032	-0.219	-0.089	
					4	0.183	-0.026	-0.018	
					Other	0.036	-0.727	-0.308	
26	0-26	0.651	0.670	0.520	1	0.100	-0.353	-0.207	
					2	0.106	-0.352	-0.209	
					3	0.651	0.670	0.520	*
					4	0.093	-0.300	-0.172	
					Other	0.050	-0.691	-0.328	
27	0-27	0.500	0.395	0.315	1	0.127	-0.269	-0.168	
					2	0.500	0.395	0.315	*
					3	0.090	-0.263	-0.149	
					4	0.238	0.010	0.007	
					Other	0.045	-0.656	-0.299	
28	0-28	0.627	0.696	0.545	1	0.116	-0.245	-0.149	
					2	0.133	-0.324	-0.205	
					3	0.091	-0.550	-0.313	
					4	0.627	0.696	0.545	*
					Other	0.032	-0.770	-0.315	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 5

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
29	0-29	0.525	0.493	0.393	1	0.525	0.493	0.393	*
					2	0.063	-0.291	-0.148	
					3	0.131	-0.217	-0.136	
					4	0.237	-0.189	-0.137	
					Other	0.045	-0.592	-0.270	
30	0-30	0.550	0.613	0.487	1	0.324	-0.299	-0.230	
					2	0.052	-0.462	-0.221	
					3	0.550	0.613	0.487	*
					4	0.023	-0.170	-0.062	
					Other	0.050	-0.740	-0.350	
31	0-31	0.597	0.296	0.234	1	0.247	0.034	0.025	
					2	0.597	0.296	0.234	*
					3	0.052	-0.041	-0.020	
					4	0.038	-0.264	-0.114	
					Other	0.066	-0.773	-0.400	
32	0-32	0.396	0.607	0.478	1	0.211	-0.012	-0.008	
					2	0.111	-0.062	-0.037	
					3	0.396	0.607	0.478	*
					4	0.203	-0.365	-0.256	
					Other	0.079	-0.786	-0.429	
33	0-33	0.622	0.417	0.327	1	0.158	0.060	0.040	
					2	0.077	-0.214	-0.116	
					3	0.622	0.417	0.327	*
					4	0.066	-0.273	-0.141	
					Other	0.077	-0.741	-0.401	
34	0-34	0.573	0.533	0.422	1	0.573	0.533	0.422	*
					2	0.090	-0.191	-0.108	
					3	0.102	-0.349	-0.205	
					4	0.152	0.016	0.010	
					Other	0.082	-0.785	-0.434	
35	0-35	0.577	0.633	0.502	1	0.129	-0.405	-0.254	
					2	0.036	-0.410	-0.174	
					3	0.179	-0.042	-0.029	
					4	0.577	0.633	0.502	*
					Other	0.079	-0.812	-0.443	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 6

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
36	0-36	0.507	0.552	0.440	1	0.032	-0.355	-0.145	
					2	0.115	-0.310	-0.188	
					3	0.507	0.552	0.440	*
					4	0.271	-0.036	-0.027	
					Other	0.075	-0.863	-0.464	
37	0-37	0.823	0.467	0.318	1	0.823	0.467	0.318	*
					2	0.050	-0.073	-0.035	
					3	0.036	0.007	0.003	
					4	0.022	-0.124	-0.044	
					Other	0.070	-0.806	-0.424	
38	0-38	0.477	0.352	0.281	1	0.025	-0.055	-0.020	
					2	0.477	0.352	0.281	*
					3	0.125	-0.032	-0.020	
					4	0.289	-0.048	-0.037	
					Other	0.084	-0.737	-0.410	
39	0-39	0.455	0.640	0.510	1	0.188	-0.292	-0.202	
					2	0.065	-0.315	-0.162	
					3	0.204	-0.038	-0.027	
					4	0.455	0.640	0.510	*
					Other	0.088	-0.781	-0.440	
40	0-40	0.337	0.649	0.501	1	0.487	-0.145	-0.116	
					2	0.045	-0.300	-0.137	
					3	0.052	-0.281	-0.135	
					4	0.337	0.649	0.501	*
					Other	0.079	-0.822	-0.448	
41	0-41	0.299	0.479	0.364	1	0.357	0.059	0.046	
					2	0.299	0.479	0.364	*
					3	0.145	0.085	0.055	
					4	0.122	-0.430	-0.266	
					Other	0.077	-0.838	-0.454	
42	0-42	0.572	0.760	0.603	1	0.122	-0.304	-0.188	
					2	0.154	-0.314	-0.206	
					3	0.572	0.760	0.603	*
					4	0.070	-0.239	-0.126	
					Other	0.082	-0.857	-0.474	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 7

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
43	0-43	0.496	0.659	0.526	1	0.143	-0.036	-0.023	
					2	0.496	0.659	0.526	*
					3	0.075	-0.223	-0.120	
					4	0.199	-0.327	-0.228	
					Other	0.086	-0.841	-0.471	
44	0-44	0.466	0.433	0.345	1	0.466	0.433	0.345	*
					2	0.203	0.016	0.011	
					3	0.151	-0.198	-0.129	
					4	0.093	0.001	0.001	
					Other	0.088	-0.819	-0.461	
45	0-45	0.548	0.664	0.529	1	0.108	-0.448	-0.267	
					2	0.548	0.664	0.529	*
					3	0.127	-0.213	-0.133	
					4	0.118	0.065	0.040	
					Other	0.099	-0.856	-0.499	
46	0-46	0.735	0.699	0.519	1	0.056	-0.211	-0.103	
					2	0.038	-0.103	-0.044	
					3	0.056	-0.171	-0.084	
					4	0.735	0.699	0.519	*
					Other	0.116	-0.909	-0.555	
47	0-47	0.631	0.631	0.493	1	0.023	-0.396	-0.145	
					2	0.097	-0.220	-0.128	
					3	0.631	0.631	0.493	*
					4	0.129	0.001	0.000	
					Other	0.120	-0.893	-0.550	
48	0-48	0.529	0.533	0.425	1	0.068	0.004	0.002	
					2	0.145	-0.110	-0.071	
					3	0.529	0.533	0.425	*
					4	0.145	-0.049	-0.031	
					Other	0.113	-0.920	-0.557	
49	0-49	0.355	0.447	0.348	1	0.145	0.087	0.056	
					2	0.297	-0.040	-0.031	
					3	0.355	0.447	0.348	*
					4	0.082	0.044	0.025	
					Other	0.120	-0.895	-0.551	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 8

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
50	0-50	0.561	0.659	0.523	1	0.097	-0.106	-0.061	
					2	0.561	0.659	0.523	*
					3	0.170	-0.094	-0.063	
					4	0.039	-0.253	-0.111	
					Other	0.133	-0.915	-0.579	
51	0-51	0.543	0.710	0.565	1	0.194	-0.121	-0.084	
					2	0.543	0.710	0.565	*
					3	0.030	-0.385	-0.154	
					4	0.108	-0.154	-0.092	
					Other	0.125	-0.937	-0.584	
52	0-52	0.683	0.875	0.670	1	0.030	-0.145	-0.058	
					2	0.683	0.875	0.670	*
					3	0.073	-0.403	-0.215	
					4	0.102	-0.401	-0.236	
					Other	0.111	-0.920	-0.554	
53	0-53	0.620	0.811	0.636	1	0.025	-0.130	-0.049	
					2	0.091	-0.201	-0.115	
					3	0.620	0.811	0.636	*
					4	0.142	-0.385	-0.247	
					Other	0.122	-0.900	-0.556	
54	0-54	0.762	0.897	0.652	1	0.061	-0.374	-0.188	
					2	0.034	-0.329	-0.137	
					3	0.023	-0.380	-0.139	
					4	0.762	0.897	0.652	*
					Other	0.120	-0.934	-0.575	
55	0-55	0.550	0.837	0.666	1	0.063	-0.306	-0.156	
					2	0.550	0.837	0.666	*
					3	0.133	-0.223	-0.141	
					4	0.129	-0.246	-0.155	
					Other	0.125	-0.940	-0.585	
56	0-56	0.670	0.811	0.625	1	0.054	-0.390	-0.189	
					2	0.056	-0.250	-0.122	
					3	0.086	-0.122	-0.068	
					4	0.670	0.811	0.625	*
					Other	0.134	-0.942	-0.598	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 9

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
57	0-57	0.507	0.575	0.459	1	0.507	0.575	0.459	*
					2	0.122	-0.062	-0.038	
					3	0.088	-0.091	-0.051	
					4	0.145	-0.013	-0.008	
					Other	0.138	-0.904	-0.578	
58	0-58	0.572	0.692	0.549	1	0.172	0.031	0.021	
					2	0.072	-0.379	-0.201	
					3	0.572	0.692	0.549	*
					4	0.047	-0.287	-0.133	
					Other	0.138	-0.906	-0.579	
59	0-59	0.668	0.857	0.661	1	0.668	0.857	0.661	*
					2	0.065	-0.231	-0.118	
					3	0.124	-0.389	-0.242	
					4	0.018	-0.328	-0.109	
					Other	0.125	-0.911	-0.568	
60	0-60	0.663	0.838	0.647	1	0.099	-0.385	-0.224	
					2	0.043	-0.148	-0.067	
					3	0.663	0.838	0.647	*
					4	0.070	-0.264	-0.139	
					Other	0.125	-0.922	-0.574	
61	0-61	0.561	0.527	0.418	1	0.032	-0.243	-0.099	
					2	0.054	-0.053	-0.026	
					3	0.561	0.527	0.418	*
					4	0.192	0.193	0.134	
					Other	0.161	-0.969	-0.644	
62	0-62	0.651	0.781	0.606	1	0.651	0.781	0.606	*
					2	0.036	-0.025	-0.011	
					3	0.120	-0.133	-0.082	
					4	0.036	-0.344	-0.146	
					Other	0.158	-0.968	-0.640	
63	0-63	0.376	0.635	0.498	1	0.206	0.050	0.035	
					2	0.131	0.063	0.040	
					3	0.376	0.635	0.498	*
					4	0.124	-0.185	-0.115	
					Other	0.163	-0.938	-0.626	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 10

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
64	0-64	0.697	0.821	0.624	1	0.697	0.821	0.624	*
					2	0.029	-0.161	-0.063	
					3	0.034	-0.337	-0.140	
					4	0.079	-0.088	-0.048	
					Other	0.161	-0.972	-0.647	
65	0-65	0.328	0.529	0.407	1	0.063	-0.093	-0.047	
					2	0.179	0.047	0.032	
					3	0.328	0.529	0.407	*
					4	0.240	0.144	0.105	
					Other	0.190	-0.872	-0.603	
66	0-66	0.410	0.678	0.536	1	0.032	-0.139	-0.057	
					2	0.056	-0.036	-0.017	
					3	0.410	0.678	0.536	*
					4	0.337	-0.043	-0.033	
					Other	0.165	-0.944	-0.631	
67	0-67	0.430	0.588	0.466	1	0.430	0.588	0.466	*
					2	0.082	-0.238	-0.132	
					3	0.056	0.007	0.003	
					4	0.265	0.145	0.107	
					Other	0.167	-0.971	-0.651	
68	0-68	0.332	0.641	0.494	1	0.186	0.290	0.199	
					2	0.267	-0.223	-0.166	
					3	0.041	0.029	0.013	
					4	0.332	0.641	0.494	*
					Other	0.174	-0.933	-0.632	
69	0-69	0.385	0.635	0.498	1	0.385	0.635	0.498	*
					2	0.036	-0.235	-0.099	
					3	0.328	0.023	0.018	
					4	0.088	0.038	0.021	
					Other	0.163	-0.968	-0.645	
70	0-70	0.529	0.741	0.591	1	0.082	-0.219	-0.121	
					2	0.190	-0.029	-0.020	
					3	0.529	0.741	0.591	*
					4	0.023	-0.234	-0.086	
					Other	0.176	-0.932	-0.633	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 11

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
71	0-71	0.317	0.643	0.492	1	0.317	0.643	0.492	*
					2	0.174	-0.014	-0.010	
					3	0.072	-0.051	-0.027	
					4	0.263	0.090	0.067	
					Other	0.174	-0.965	-0.654	
72	0-72	0.446	0.564	0.449	1	0.018	-0.370	-0.123	
					2	0.446	0.564	0.449	*
					3	0.023	-0.218	-0.080	
					4	0.351	0.113	0.088	
					Other	0.161	-0.967	-0.643	
73	0-73	0.625	0.773	0.605	1	0.057	-0.129	-0.064	
					2	0.625	0.773	0.605	*
					3	0.065	-0.115	-0.059	
					4	0.084	-0.118	-0.066	
					Other	0.168	-0.976	-0.656	
74	0-74	0.643	0.684	0.533	1	0.643	0.684	0.533	*
					2	0.095	0.135	0.078	
					3	0.032	-0.173	-0.071	
					4	0.063	-0.231	-0.118	
					Other	0.167	-0.948	-0.636	
75	0-75	0.656	0.839	0.650	1	0.081	-0.182	-0.100	
					2	0.656	0.839	0.650	*
					3	0.048	-0.150	-0.070	
					4	0.041	-0.188	-0.084	
					Other	0.174	-0.973	-0.659	
76	0-76	0.504	0.757	0.604	1	0.504	0.757	0.604	*
					2	0.192	-0.044	-0.031	
					3	0.077	-0.053	-0.029	
					4	0.034	-0.145	-0.060	
					Other	0.194	-0.988	-0.687	
77	0-77	0.489	0.809	0.646	1	0.082	-0.135	-0.075	
					2	0.489	0.809	0.646	*
					3	0.102	-0.150	-0.088	
					4	0.125	0.005	0.003	
					Other	0.201	-0.985	-0.690	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 12

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
78	0-78	0.613	0.682	0.536	1	0.613	0.682	0.536	*
					2	0.125	0.178	0.111	
					3	0.036	-0.098	-0.042	
					4	0.032	-0.205	-0.084	
					Other	0.194	-1.000	-0.696	
79	0-79	0.523	0.840	0.669	1	0.073	0.011	0.006	
					2	0.133	-0.175	-0.111	
					3	0.065	-0.133	-0.068	
					4	0.523	0.840	0.669	*
					Other	0.206	-0.988	-0.696	
80	0-80	0.355	0.688	0.535	1	0.147	0.070	0.045	
					2	0.247	0.071	0.052	
					3	0.054	-0.171	-0.083	
					4	0.355	0.688	0.535	*
					Other	0.197	-0.994	-0.693	
81	0-81	0.419	0.468	0.370	1	0.306	0.328	0.250	
					2	0.052	-0.093	-0.045	
					3	0.419	0.468	0.370	*
					4	0.016	-0.023	-0.008	
					Other	0.206	-1.000	-0.710	
82	0-82	0.672	0.885	0.681	1	0.047	-0.061	-0.028	
					2	0.672	0.885	0.681	*
					3	0.048	-0.186	-0.087	
					4	0.034	-0.196	-0.082	
					Other	0.199	-1.000	-0.702	
83	0-83	0.441	0.600	0.477	1	0.163	0.215	0.144	
					2	0.441	0.600	0.477	*
					3	0.084	0.058	0.032	
					4	0.100	-0.077	-0.045	
					Other	0.211	-0.986	-0.699	
84	0-84	0.385	0.668	0.525	1	0.238	0.136	0.099	
					2	0.136	-0.034	-0.021	
					3	0.385	0.668	0.525	*
					4	0.030	-0.011	-0.005	
					Other	0.210	-1.000	-0.711	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 13

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
85	0-85	0.364	0.618	0.482	1	0.084	-0.118	-0.066	
					2	0.151	0.246	0.161	
					3	0.364	0.618	0.482	*
					4	0.188	0.057	0.039	
					Other	0.213	-0.986	-0.700	
86	0-86	0.541	0.787	0.627	1	0.541	0.787	0.627	*
					2	0.134	-0.021	-0.013	
					3	0.072	-0.009	-0.005	
					4	0.041	-0.192	-0.085	
					Other	0.211	-1.000	-0.709	
87	0-87	0.514	0.806	0.643	1	0.032	-0.190	-0.078	
					2	0.066	-0.046	-0.024	
					3	0.514	0.806	0.643	*
					4	0.181	-0.054	-0.037	
					Other	0.206	-1.000	-0.710	
88	0-88	0.262	0.555	0.411	1	0.262	0.555	0.411	*
					2	0.172	0.165	0.111	
					3	0.147	0.097	0.063	
					4	0.194	0.197	0.137	
					Other	0.226	-0.995	-0.715	
89	0-89	0.319	0.609	0.466	1	0.165	0.079	0.053	
					2	0.077	0.038	0.020	
					3	0.319	0.609	0.466	*
					4	0.213	0.182	0.130	
					Other	0.226	-0.984	-0.707	
90	0-90	0.353	0.674	0.524	1	0.211	0.065	0.046	
					2	0.116	-0.069	-0.042	
					3	0.097	0.235	0.136	
					4	0.353	0.674	0.524	*
					Other	0.222	-0.994	-0.712	
91	0-91	0.303	0.645	0.490	1	0.204	0.148	0.104	
					2	0.303	0.645	0.490	*
					3	0.152	0.158	0.104	
					4	0.068	0.236	0.123	
					Other	0.272	-1.000	-0.753	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 14

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
92	0-92	0.487	0.885	0.706	1	0.215	-0.054	-0.039	
					2	0.029	0.025	0.010	
					3	0.487	0.885	0.706	*
					4	0.005	-0.196	-0.041	
					Other	0.263	-1.000	-0.762	
93	0-93	0.358	0.602	0.469	1	0.358	0.602	0.469	*
					2	0.170	0.157	0.106	
					3	0.152	0.218	0.143	
					4	0.050	0.220	0.104	
					Other	0.269	-1.000	-0.764	
94	0-94	0.473	0.792	0.632	1	0.059	0.014	0.007	
					2	0.065	-0.011	-0.006	
					3	0.129	0.115	0.072	
					4	0.473	0.792	0.632	*
					Other	0.274	-1.000	-0.762	
95	0-95	0.305	0.661	0.503	1	0.106	0.204	0.121	
					2	0.276	0.215	0.161	
					3	0.305	0.661	0.503	*
					4	0.043	-0.044	-0.020	
					Other	0.271	-1.000	-0.757	
96	0-96	0.450	0.785	0.625	1	0.056	0.189	0.093	
					2	0.102	-0.122	-0.072	
					3	0.450	0.785	0.625	*
					4	0.124	0.138	0.085	
					Other	0.269	-1.000	-0.763	
97	0-97	0.428	0.730	0.579	1	0.428	0.730	0.579	*
					2	0.163	0.258	0.172	
					3	0.070	0.014	0.007	
					4	0.066	-0.132	-0.068	
					Other	0.272	-1.000	-0.752	
98	0-98	0.398	0.822	0.648	1	0.398	0.822	0.648	*
					2	0.077	-0.024	-0.013	
					3	0.104	-0.025	-0.015	
					4	0.125	0.171	0.107	
					Other	0.296	-0.998	-0.755	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 15

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
99	0-99	0.452	0.818	0.651	1	0.452	0.818	0.651	*
					2	0.081	0.248	0.136	
					3	0.106	-0.034	-0.020	
					4	0.084	-0.089	-0.050	
					Other	0.278	-1.000	-0.761	
100	0-100	0.554	0.811	0.645	1	0.041	0.046	0.021	
					2	0.554	0.811	0.645	*
					3	0.048	-0.005	-0.002	
					4	0.082	0.104	0.057	
					Other	0.274	-1.000	-0.762	
101	0-101	0.369	0.771	0.603	1	0.104	0.098	0.058	
					2	0.369	0.771	0.603	*
					3	0.097	0.025	0.014	
					4	0.147	0.129	0.084	
					Other	0.283	-1.000	-0.761	
102	0-102	0.487	0.811	0.647	1	0.075	-0.037	-0.020	
					2	0.093	0.004	0.002	
					3	0.066	0.186	0.096	
					4	0.487	0.811	0.647	*
					Other	0.278	-1.000	-0.765	
103	0-103	0.276	0.670	0.501	1	0.081	0.072	0.040	
					2	0.127	0.185	0.115	
					3	0.276	0.670	0.501	*
					4	0.231	0.234	0.169	
					Other	0.285	-1.000	-0.763	
104	0-104	0.262	0.769	0.569	1	0.262	0.769	0.569	*
					2	0.138	0.019	0.012	
					3	0.203	0.259	0.181	
					4	0.120	0.076	0.047	
					Other	0.278	-1.000	-0.765	
105	0-105	0.349	0.561	0.435	1	0.229	0.241	0.174	
					2	0.129	0.306	0.192	
					3	0.349	0.561	0.435	*
					4	0.011	-0.128	-0.035	
					Other	0.281	-1.000	-0.759	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 16

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
106	0-106	0.489	0.857	0.684	1	0.122	0.026	0.016	
					2	0.030	-0.006	-0.002	
					3	0.052	0.063	0.030	
					4	0.489	0.857	0.684	*
					Other	0.306	-1.000	-0.767	
107	0-107	0.409	0.746	0.589	1	0.097	0.177	0.102	
					2	0.409	0.746	0.589	*
					3	0.095	0.136	0.078	
					4	0.095	0.081	0.047	
					Other	0.305	-1.000	-0.775	
108	0-108	0.324	0.686	0.527	1	0.324	0.686	0.527	*
					2	0.084	0.039	0.022	
					3	0.102	0.294	0.173	
					4	0.176	0.179	0.122	
					Other	0.314	-0.991	-0.758	
109	0-109	0.323	0.568	0.436	1	0.009	0.196	0.050	
					2	0.323	0.568	0.436	*
					3	0.122	0.004	0.002	
					4	0.238	0.483	0.351	
					Other	0.308	-1.000	-0.778	
110	0-110	0.285	0.681	0.512	1	0.158	0.215	0.142	
					2	0.081	0.021	0.011	
					3	0.285	0.681	0.512	*
					4	0.163	0.291	0.194	
					Other	0.314	-1.000	-0.771	
111	0-111	0.498	0.929	0.741	1	0.059	-0.086	-0.043	
					2	0.498	0.929	0.741	*
					3	0.027	-0.133	-0.051	
					4	0.100	0.039	0.023	
					Other	0.315	-1.000	-0.773	
112	0-112	0.351	0.707	0.549	1	0.351	0.707	0.549	*
					2	0.054	-0.050	-0.024	
					3	0.118	0.133	0.082	
					4	0.159	0.304	0.202	
					Other	0.317	-1.000	-0.767	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 17

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
113	0-113	0.246	0.724	0.529	1	0.043	-0.142	-0.064	
					2	0.246	0.724	0.529	*
					3	0.016	-0.039	-0.013	
					4	0.376	0.395	0.310	
					Other	0.319	-1.000	-0.780	
114	0-114	0.317	0.589	0.451	1	0.088	0.120	0.068	
					2	0.158	0.384	0.254	
					3	0.317	0.589	0.451	*
					4	0.104	0.224	0.132	
					Other	0.333	-0.996	-0.768	
115	0-115	0.405	0.844	0.667	1	0.405	0.844	0.667	*
					2	0.061	-0.116	-0.058	
					3	0.099	0.074	0.043	
					4	0.113	0.184	0.111	
					Other	0.323	-1.000	-0.773	
116	0-116	0.373	0.665	0.521	1	0.129	0.215	0.135	
					2	0.038	0.120	0.052	
					3	0.373	0.665	0.521	*
					4	0.136	0.244	0.155	
					Other	0.324	-1.000	-0.769	
117	0-117	0.265	0.701	0.520	1	0.254	0.302	0.222	
					2	0.043	0.230	0.104	
					3	0.108	0.056	0.033	
					4	0.265	0.701	0.520	*
					Other	0.330	-0.988	-0.761	
118	0-118	0.285	0.628	0.472	1	0.285	0.628	0.472	*
					2	0.289	0.374	0.282	
					3	0.059	0.020	0.010	
					4	0.043	0.209	0.094	
					Other	0.324	-1.000	-0.774	
119	0-119	0.412	0.755	0.597	1	0.142	0.221	0.142	
					2	0.041	-0.061	-0.027	
					3	0.412	0.755	0.597	*
					4	0.068	0.216	0.113	
					Other	0.337	-1.000	-0.775	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 18

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
120	0-120	0.486	0.805	0.642	1	0.081	0.054	0.030	
					2	0.043	0.021	0.010	
					3	0.059	0.266	0.133	
					4	0.486	0.805	0.642	*
					Other	0.332	-0.999	-0.770	
121	0-121	0.269	0.603	0.449	1	0.242	0.316	0.230	
					2	0.269	0.603	0.449	*
					3	0.041	0.285	0.126	
					4	0.068	0.397	0.207	
					Other	0.380	-0.984	-0.772	
122	0-122	0.341	0.642	0.497	1	0.341	0.642	0.497	*
					2	0.029	0.069	0.027	
					3	0.163	0.465	0.310	
					4	0.086	0.180	0.101	
					Other	0.382	-1.000	-0.788	
123	0-123	0.301	0.723	0.549	1	0.050	0.014	0.007	
					2	0.213	0.340	0.242	
					3	0.301	0.723	0.549	*
					4	0.054	0.251	0.122	
					Other	0.382	-0.995	-0.781	
124	0-124	0.382	0.822	0.645	1	0.070	0.172	0.090	
					2	0.066	0.136	0.070	
					3	0.091	0.138	0.079	
					4	0.382	0.822	0.645	*
					Other	0.391	-0.981	-0.772	
125	0-125	0.389	0.844	0.664	1	0.030	-0.063	-0.025	
					2	0.118	0.208	0.128	
					3	0.389	0.844	0.664	*
					4	0.063	0.186	0.095	
					Other	0.400	-0.992	-0.783	
126	0-126	0.242	0.672	0.490	1	0.138	0.379	0.242	
					2	0.242	0.672	0.490	*
					3	0.090	0.144	0.082	
					4	0.120	0.343	0.211	
					Other	0.410	-0.991	-0.783	

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 19

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
127	0-127	0.310	0.807	0.616	1	0.310	0.807	0.616	*
					2	0.104	0.185	0.109	
					3	0.081	0.131	0.072	
					4	0.099	0.285	0.166	
					Other	0.407	-0.998	-0.788	
128	0-128	0.190	0.614	0.425	1	0.249	0.323	0.237	
					2	0.190	0.614	0.425	*
					3	0.099	0.432	0.252	
					4	0.057	0.366	0.181	
					Other	0.405	-0.996	-0.786	
129	0-129	0.362	0.858	0.669	1	0.362	0.858	0.669	*
					2	0.047	0.131	0.061	
					3	0.145	0.185	0.120	
					4	0.041	0.107	0.047	
					Other	0.405	-0.996	-0.786	
130	0-130	0.360	0.890	0.693	1	0.063	0.029	0.015	
					2	0.100	0.221	0.129	
					3	0.070	0.080	0.042	
					4	0.360	0.890	0.693	*
					Other	0.407	-0.995	-0.786	
131	0-131	0.208	0.774	0.547	1	0.208	0.774	0.547	*
					2	0.038	0.226	0.097	
					3	0.086	0.213	0.120	
					4	0.258	0.346	0.255	
					Other	0.410	-0.992	-0.784	
132	0-132	0.263	0.647	0.480	1	0.036	0.132	0.056	
					2	0.246	0.442	0.323	
					3	0.263	0.647	0.480	*
					4	0.041	0.285	0.126	
					Other	0.414	-0.990	-0.783	
133	0-133	0.297	0.535	0.405	1	0.240	0.507	0.369	
					2	0.297	0.535	0.405	*
					3	0.048	0.383	0.179	
					4	0.004	0.008	0.001	
					Other	0.410	-0.981	-0.775	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 20

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				Key
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	
134	0-134	0.482	0.876	0.699	1	0.025	0.073	0.027	*
					2	0.009	0.061	0.016	
					3	0.482	0.876	0.699	
					4	0.079	0.205	0.112	
					Other	0.405	-0.994	-0.785	
135	0-135	0.213	0.651	0.462	1	0.213	0.651	0.462	*
					2	0.099	0.115	0.067	
					3	0.251	0.503	0.370	
					4	0.022	0.292	0.104	
					Other	0.416	-0.986	-0.781	
136	0-136	0.450	0.875	0.696	1	0.057	0.208	0.103	*
					2	0.450	0.875	0.696	
					3	0.030	0.087	0.035	
					4	0.036	0.043	0.018	
					Other	0.427	-0.968	-0.767	
137	0-137	0.274	0.799	0.597	1	0.099	0.183	0.107	*
					2	0.124	0.342	0.212	
					3	0.274	0.799	0.597	
					4	0.072	0.108	0.057	
					Other	0.432	-0.974	-0.773	
138	0-138	0.203	0.626	0.439	1	0.100	0.157	0.092	*
					2	0.203	0.626	0.439	
					3	0.167	0.465	0.312	
					4	0.100	0.339	0.198	
					Other	0.430	-0.967	-0.767	
139	0-139	0.242	0.787	0.574	1	0.242	0.787	0.574	*
					2	0.226	0.288	0.207	
					3	0.023	0.042	0.015	
					4	0.077	0.288	0.156	
					Other	0.432	-0.957	-0.760	
140	0-140	0.167	0.662	0.444	1	0.167	0.662	0.444	*
					2	0.168	0.363	0.244	
					3	0.156	0.432	0.285	
					4	0.065	0.132	0.068	
					Other	0.444	-0.954	-0.759	

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 21

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key				
141	0-141	0.323	0.792	0.608	1	0.122	0.284	0.175					
					2	0.057	0.194	0.096					
					3	0.323	0.792	0.608	*				
					4	0.057	0.116	0.057					
					Other	0.441	-0.956	-0.760					
142	0-142	0.285	0.818	0.615	1	0.099	0.311	0.181					
					2	0.090	0.124	0.070					
					3	0.285	0.818	0.615	*				
					4	0.084	0.173	0.097					
					Other	0.443	-0.959	-0.762					
143	0-143	0.149	0.678	0.442	1	0.149	0.269	0.175					
					2	0.122	0.305	0.189					
					3	0.122	0.411	0.254					
					4	0.149	0.678	0.442	*				
					Other	0.459	-0.919	-0.732					
144	0-144	0.136	0.596	0.379	1	0.061	0.233	0.117					
					2	0.136	0.596	0.379	*				
					CHECK THE KEY				3	0.251	0.542	0.398	?
					2 was specified, 3 works better				4	0.100	0.254	0.149	
					Other	0.452	-0.948	-0.755					
145	0-145	0.220	0.714	0.510	1	0.201	0.426	0.299					
					2	0.220	0.714	0.510	*				
					3	0.045	0.142	0.065					
					4	0.088	0.196	0.110					
					Other	0.446	-0.951	-0.756					
146	0-146	0.188	0.694	0.479	1	0.167	0.377	0.253					
					2	0.188	0.694	0.479	*				
					3	0.102	0.274	0.161					
					4	0.086	0.219	0.122					
					Other	0.457	-0.919	-0.732					
147	0-147	0.353	0.839	0.652	1	0.081	0.077	0.042					
					2	0.061	0.343	0.173					
					3	0.057	0.081	0.040					
					4	0.353	0.839	0.652	*				
					Other	0.448	-0.945	-0.752					

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 22

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics							
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key			
148	0-148	0.294	0.752	0.568	1	0.129	0.334	0.210				
					2	0.100	0.171	0.100				
					3	0.294	0.752	0.568	*			
					4	0.032	0.227	0.093				
					Other	0.444	-0.951	-0.756				
149	0-149	0.235	0.793	0.575	1	0.106	0.168	0.100				
					2	0.235	0.793	0.575	*			
					3	0.048	0.153	0.072				
					4	0.161	0.339	0.225				
					Other	0.450	-0.942	-0.749				
150	0-150	0.281	0.799	0.599	1	0.281	0.799	0.599	*			
					2	0.097	0.123	0.072				
					3	0.102	0.355	0.209				
					4	0.070	0.177	0.093				
					Other	0.450	-0.954	-0.759				
151	0-151	0.066	0.619	0.320	1	0.349	0.662	0.514	?			
					2	0.032	0.181	0.074				
					CHECK THE KEY			3	0.009	-0.018	-0.005	
					4 was specified, 1 works better			4	0.066	0.619	0.320	*
								Other	0.543	-0.851	-0.677	
152	0-152	0.183	0.684	0.469	1	0.183	0.684	0.469	*			
					2	0.118	0.246	0.151				
					3	0.104	0.482	0.285				
					4	0.045	0.226	0.103				
					Other	0.550	-0.854	-0.680				
153	0-153	0.102	0.522	0.307	1	0.088	0.428	0.241				
					2	0.102	0.522	0.307	*			
					CHECK THE KEY			3	0.081	0.328	0.180	
					2 was specified, 4 works better			4	0.165	0.511	0.342	?
								Other	0.565	-0.856	-0.680	
154	0-154	0.145	0.643	0.416	1	0.079	0.317	0.173				
					2	0.072	0.361	0.191				
					3	0.145	0.643	0.416	*			
					4	0.134	0.396	0.251				
					Other	0.570	-0.836	-0.663				

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 23

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key				
155	0-155	0.124	0.477	0.296	1	0.276	0.688	0.514	?				
					2	0.022	0.038	0.014					
					CHECK THE KEY				3	0.016	0.085	0.027	
					4 was specified, 1 works better				4	0.124	0.477	0.296	*
					Other	0.563	-0.844	-0.670					
156	0-156	0.222	0.559	0.401	1	0.222	0.559	0.401	*				
					2	0.048	0.196	0.092					
					3	0.061	0.300	0.151					
					4	0.102	0.608	0.358					
					Other	0.566	-0.841	-0.668					
157	0-157	0.224	0.672	0.482	1	0.086	0.430	0.241					
					2	0.034	0.061	0.026					
					3	0.084	0.350	0.195					
					4	0.224	0.672	0.482	*				
					Other	0.572	-0.834	-0.661					
158	0-158	0.041	0.327	0.145	1	0.041	0.327	0.145	*				
					2	0.115	0.465	0.282					
					CHECK THE KEY				3	0.211	0.598	0.424	?
					1 was specified, 3 works better				4	0.054	0.272	0.132	
					Other	0.579	-0.823	-0.652					
159	0-159	0.070	0.531	0.279	1	0.215	0.463	0.329	?				
					2	0.050	0.374	0.177					
					CHECK THE KEY				3	0.086	0.507	0.284	
					4 was specified, 1 works better				4	0.070	0.531	0.279	*
					Other	0.579	-0.830	-0.657					
160	0-160	0.201	0.746	0.523	1	0.201	0.746	0.523	*				
					2	0.029	0.100	0.039					
					3	0.174	0.359	0.243					
					4	0.016	0.269	0.086					
					Other	0.581	-0.816	-0.646					
161	0-161	0.210	0.725	0.512	1	0.210	0.725	0.512	*				
					2	0.045	0.217	0.099					
					3	0.124	0.334	0.207					
					4	0.041	0.246	0.109					
					Other	0.581	-0.816	-0.646					

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 24

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key				
162	0-162	0.254	0.706	0.520	1	0.030	0.178	0.071					
					2	0.254	0.706	0.520	*				
					3	0.018	0.059	0.020					
					4	0.106	0.395	0.235					
					Other	0.591	-0.807	-0.638					
163	0-163	0.151	0.635	0.415	1	0.047	0.283	0.131					
					2	0.151	0.635	0.415	*				
					3	0.131	0.418	0.263					
					4	0.068	0.310	0.162					
					Other	0.604	-0.793	-0.625					
164	0-164	0.267	0.793	0.589	1	0.267	0.793	0.589	*				
					2	0.068	0.140	0.073					
					3	0.025	0.057	0.021					
					4	0.045	0.304	0.139					
					Other	0.595	-0.803	-0.634					
165	0-165	0.106	0.617	0.367	1	0.106	0.617	0.367	*				
					2	0.091	0.331	0.189					
					3	0.183	0.451	0.309					
					4	0.027	0.381	0.146					
					Other	0.593	-0.800	-0.632					
166	0-166	0.043	0.157	0.071	1	0.065	0.446	0.229					
					2	0.043	0.157	0.071	*				
					CHECK THE KEY				3	0.194	0.602	0.418	?
					2 was specified, 3 works better				4	0.086	0.381	0.213	
					Other	0.613	-0.772	-0.607					
167	0-167	0.108	0.642	0.383	1	0.124	0.285	0.177					
					2	0.095	0.499	0.287					
					3	0.108	0.642	0.383	*				
					4	0.059	0.297	0.148					
					Other	0.615	-0.775	-0.609					
168	0-168	0.206	0.708	0.499	1	0.206	0.708	0.499	*				
					2	0.061	0.156	0.079					
					3	0.097	0.394	0.228					
					4	0.014	0.093	0.029					
					Other	0.622	-0.766	-0.601					

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 25

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics							
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key			
169	0-169	0.194	0.687	0.477	1	0.194	0.687	0.477	*			
					2	0.050	0.199	0.094				
					3	0.075	0.339	0.182				
					4	0.052	0.315	0.151				
					Other	0.629	-0.769	-0.602				
170	0-170	0.179	0.588	0.401	1	0.052	0.402	0.192				
					2	0.066	0.240	0.124				
					3	0.179	0.588	0.401	*			
					4	0.075	0.463	0.249				
					Other	0.627	-0.775	-0.607				
171	0-171	0.057	0.340	0.168	1	0.190	0.652	0.451	?			
					2	0.057	0.340	0.168	*			
					CHECK THE KEY			3	0.023	0.216	0.079	
					2 was specified, 1 works better			4	0.100	0.379	0.222	
					Other	0.629	-0.780	-0.610				
172	0-172	0.038	0.487	0.210	1	0.043	0.060	0.027				
					2	0.030	0.272	0.109				
					CHECK THE KEY			3	0.038	0.487	0.210	*
					3 was specified, 4 works better			4	0.254	0.714	0.526	?
					Other	0.634	-0.780	-0.609				
173	0-173	0.100	0.600	0.351	1	0.151	0.467	0.305				
					2	0.036	0.270	0.115				
					3	0.063	0.361	0.183				
					4	0.100	0.600	0.351	*			
					Other	0.651	-0.758	-0.588				
174	0-174	0.061	0.563	0.284	1	0.061	0.563	0.284	*			
					2	0.154	0.411	0.270				
					CHECK THE KEY			3	0.118	0.545	0.334	?
					1 was specified, 3 works better			4	0.029	0.260	0.102	
					Other	0.638	-0.774	-0.604				
175	0-175	0.057	0.429	0.212	1	0.057	0.429	0.212	*			
					2	0.194	0.657	0.456	?			
					CHECK THE KEY			3	0.086	0.268	0.150	
					1 was specified, 2 works better			4	0.025	0.248	0.093	
					Other	0.638	-0.764	-0.596				

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 26

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key				
176	0-176	0.170	0.611	0.411	1	0.066	0.257	0.133					
					2	0.170	0.611	0.411	*				
					3	0.066	0.467	0.242					
					4	0.059	0.339	0.169					
					Other	0.638	-0.768	-0.599					
177	0-177	0.057	0.364	0.180	1	0.048	0.206	0.097					
					2	0.057	0.364	0.180	*				
					CHECK THE KEY				3	0.022	0.174	0.062	
					2 was specified, 4 works better				4	0.231	0.706	0.510	?
					Other	0.642	-0.767	-0.597					
178	0-178	0.120	0.575	0.354	1	0.104	0.414	0.245					
					2	0.090	0.441	0.250					
					3	0.120	0.575	0.354	*				
					4	0.038	0.249	0.107					
					Other	0.649	-0.759	-0.590					
179	0-179	0.034	0.533	0.222	1	0.027	0.332	0.128					
					2	0.034	0.533	0.222	*				
					CHECK THE KEY				3	0.256	0.611	0.450	?
					2 was specified, 3 works better				4	0.032	0.312	0.128	
					Other	0.651	-0.757	-0.588					
180	0-180	0.066	0.511	0.264	1	0.095	0.419	0.241					
					2	0.115	0.321	0.195					
					CHECK THE KEY				3	0.086	0.557	0.312	?
					4 was specified, 3 works better				4	0.066	0.511	0.264	*
					Other	0.638	-0.764	-0.596					

MicroCAT (tm) Testing System

Copyright (c) 1982, 1984, 1986 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN Version 2.01

Item analysis for data from file lenguas.dat

Page 27

There were 558 examinees in the data file.

Scale Statistics

```

-----
Scale:          0
-----
N of Items      180
N of Examinees 558
Mean            74.536
Variance        1554.227
Std. Dev.       39.424
Skew            -0.038
Kurtosis        -0.981
Minimum         1.000
Maximum         162.000
Median          81.000
Alpha           0.982
SEM             5.259
Mean P          0.414
Mean Item-Tot. 0.482
Mean Biserial   0.647

```

→age 27

There were 558 examinees in the data file.

Scale Statistics
