



**Universidad Autónoma de Querétaro  
Facultad de Lenguas y Letras**

**COMPARACIÓN DE TOEFL Y FIRST CERTIFICATE:  
VALIDEZ Y CONFIABILIDAD**

**TESIS**

**QUE COMO PARTE DE LOS REQUISITOS PARA OBTENER EL  
GRADO DE**

**MAESTRÍA EN LINGÜÍSTICA**

**PRESENTA**

**ESTELA MENDOZA GONZÁLEZ**

**Santiago de Querétaro, abril, 2009.**



**Universidad Autónoma de Querétaro  
Facultad de Lenguas y Letras**

**MAESTRÍA EN LINGÜÍSTICA**

**“COMPARACIÓN DE TOEFL Y FIRST CERTIFICATE: VALIDEZ Y  
CONFIABILIDAD”**

TESIS

QUE COMO PARTE DE LOS REQUISITOS PARA OBTENER EL GRADO DE  
MAESTRA EN LINGÜÍSTICA

PRESENTA  
**ESTELA MENDOZA GONZÁLEZ**

DIRIGIDO POR:

**DRA. DONNA JACKSON LEMBARK**

SINODALES

**DRA. DONNA JACKSON LEMBARK**  
Presidente

**DR. IGNACIO RODRÍGUEZ SÁNCHEZ**  
Secretario

**DRA. LUISA JOSEFINA ALARCÓN NEVE**  
Vocal

**MTRA. PHYLLIS HERRIN DE OBREGÓN**  
Suplente

**MTRA. ROSA PATRICA BÁRCENAS ACOSTA**  
Suplente

**LLM-E MA. EUGENIA CASTILLEJOS SOLIS**  
Directora Interina de la Facultad de  
Lenguas y Letras

**DR. LUIS GERARDO HERNÁNDEZ SANDOVAL**  
Director de Investigación y Posgrado

CENTRO UNIVERSITARIO  
QUERÉTARO, QRO.  
ABRIL, 2009  
MÉXICO

## RESUMEN

Esta investigación se llevó a cabo con un grupo de adolescentes mexicanos, alumnos de nivel de inglés avanzado, con el objetivo de comparar los puntajes obtenidos en dos pruebas internacionales de proficiencia en inglés, la Test of English as a Foreign Language (TOEFL) del Educational Testing Service (ETS) de Estados Unidos, y el First Certificate in English (FCE) de la Universidad de Cambridge, Inglaterra. Las hipótesis centrales sostienen primeramente, que los participantes tendrán un mejor desempeño en la prueba FCE y que por medio de los puntajes obtenidos se corroborarán los hallazgos reportados por algunos investigadores con respecto a la consistencia interna de estas pruebas, donde afirman que la consistencia interna del TOEFL es más fuerte. También se postula que el FCE tiene una validez con una alta significancia a diferencia del TOEFL. Participaron 23 sujetos que tomaban preparación para presentar el examen FCE y que no conocían el examen TOEFL. Se aplicó primeramente el examen FCE, pero solamente cuatro de las cinco subpruebas que la componen, y una semana más tarde se administró el TOEFL, pero sólo tres de las cinco subpruebas, observando los lineamientos para aplicación de dichos exámenes. En la sección del examen de Expresión Oral del FCE únicamente participaron diez sujetos, los cuales fueron entrevistados por examinadores autorizados por la Universidad de Cambridge. Debido a que se carecía del sistema de calificación oficial para estas pruebas, se tomaron los puntajes obtenidos para llevar a cabo los estadísticos necesarios para analizar las puntuaciones. Los resultados corroboran que la prueba TOEFL tiene, en las tres subpruebas, una consistencia interna más fuerte que el FCE. También se encontró que el TOEFL tiene una validez con una alta significancia en todas sus subpruebas, a diferencia del FCE donde sólo fue parcial. Finalmente los puntajes obtenidos por los participantes en la prueba FCE, demuestran su desarrollo de la proficiencia en una segunda lengua

(Palabras clave: **proficiencia, confiabilidad, consistencia interna, validez concurrente, significancia**)

## SUMMARY

This investigation was carried out with Mexican adolescents who studied English at an advanced level. The objective of this study was to compare the scores the participants got in two international proficiency tests of English, the Test of English as a Foreign Language (TOEFL) from the Educational Testing Service (ETS) from the U.S.A., and the First Certificate in English from the University of Cambridge. The central hypotheses intend first to corroborate, through their scores, that the participants will have a better performance in the FCE test and also that the internal consistency of TOEFL is stronger than the FCE's according to what has been stated by some researchers. Second, that the FCE exam has a higher significance than TOEFL. There were 23 participants, who were taking preparation to take the FCE exam and who were not familiar with TOEFL. The exams were administered following the regulations for each one of them. The first one to be taken was FCE, but it consisted of four, instead of its five parts. Then a week after, it was the turn of the TOEFL exam which consisted of only three parts, out of four. The Oral Exam in FCE was administered by examiners authorized by the University of Cambridge but it was taken by just ten subjects. Due to the lack of official scales and grading systems for the tests, the raw scores were used to process the necessary statistics to analyze the data. The results corroborate that the TOEFL exam has a stronger internal consistency than the FCE. The results also show that TOEFL has validity with a high significance in all three parts, while in the case of FCE it was partial. Finally, the scores the participants got in the TOEFL exam show the development of their proficiency in a second language.

(Key words: **proficiency, reliability, internal consistency, concurrent validity, significance**)

Mi agradecimiento a:

- Profra. Margarita Trejo Casillas, directora de inglés del Liceo Consuelo Rubio de Ruiz, por permitirme realizar esta investigación en la sección preparatoria y por su apoyo tanto en lo laboral como personal.
- Mis amigos, por su lealtad y alentarme en los momentos difíciles.
- Mis profesores de la maestría ya que todos y cada uno de ellos me hicieron crecer como persona y profesional de la enseñanza, pero particularmente a Donna Jackson-Maldonado, gracias por los cuestionamientos, a Luisa Josefina Alarcón e Ignacio Rodríguez, les agradezco sus valiables comentarios a este trabajo.

A Rodrigo Alonso  
(sorry for the unplayed games and untold stories)

A Char  
(thanks for being there, thanks for listening)

A Oshi & Meggy

## INDICE GENERAL

	Página
<b>RESUMEN</b>	<b>i</b>
<b>SUMMARY</b>	<b>ii</b>
<b>DEDICATORIA</b>	<b>iii</b>
<b>AGRADECIMIENTOS</b>	<b>iv</b>
<b>ÍNDICE</b>	<b>v</b>
<b>LISTA DE CUADROS</b>	<b>vii</b>
<b>LISTA DE TABLAS</b>	<b>viii</b>
<b>CAPÍTULO 1. INTRODUCCIÓN</b>	<b>1</b>
1.1 Objetivos	<b>1</b>
1.2 Justificación	<b>2</b>
1.3 Hipótesis	<b>5</b>
<b>CAPÍTULO 2. FUNDAMENTOS DE PRUEBAS DE PROFICIENCIA DE LENGUA</b>	<b>6</b>
2.1 Características y Tipos de Pruebas de Proficiencia de Lengua	<b>7</b>
2.2 Tipos de validez	<b>10</b>
2.3 Validez Aparente	<b>13</b>
2.4 Validez de Contenido	<b>13</b>
2.5 Validez de Constructo	<b>15</b>
2.5.1 Validez Convergente	<b>17</b>
2.5.2 Validez Discriminante	<b>18</b>
2.6 Validez Relativa al Criterio	<b>18</b>
2.6.1 Validez Predictiva	<b>20</b>
2.6.2 Validez Concurrente	<b>21</b>
2.7 Confiabilidad	<b>23</b>
2.8 Confiabilidad Interjuez o entre Evaluadores	<b>24</b>
2.9 Estimaciones de Confiabilidad de Prueba y Posprueba	<b>25</b>
2.10 Confiabilidad Formatos Paralelos	<b>25</b>
2.11 Confiabilidad: Consistencia Interna	<b>26</b>
<b>CAPÍTULO 3. PRUEBAS DE PROFICIENCIA DE CONOCIMIENTOS DE LENGUA</b>	<b>29</b>
3.1 Definición de Proficiencia de Lengua	<b>30</b>
3.2 Diferentes Tipos de Pruebas de Conocimientos de Lengua	<b>40</b>

<b>CAPÍTULO 4. FCE Y TOEFL – PRUEBAS INTERNACIONALES DE PROFICIENCIA DE LENGUA</b>	<b>50</b>
4.1 Fundamentos Teóricos del FCE	51
4.2 Sistema de Evaluación del FCE	55
4.3 Elaboración de los Componentes del Examen	56
4.3.1 Características de la Prueba de Comprensión de Lectura	61
4.3.2 Características de la Prueba de Escritura	62
4.3.3 Características la Prueba de Gramática y Vocabulario	64
4.3.4 Características de la Prueba de Comprensión Auditiva	65
4.3.5 Características de la Prueba de Expresión Oral	66
4.4 Validez y Confiabilidad en el FCE	69
4.5 Fundamentos Teóricos del TOEFL	74
4.5.1 Sistema de Evaluación del TOEFL	77
4.5.2 Elaboración de los Componentes del Examen	78
4.5.3 Características de la Prueba de Comprensión Auditiva	80
4.5.4 Características de la Prueba de Expresión Escrita y Gramática	81
4.5.5 Características de la Prueba de Vocabulario y Comprensión de Lectura	81
4.5.6 Características de la Prueba Escrita (Ensayo)	82
4.6 Validez y Confiabilidad en el TOEFL	83
<b>CAPÍTULO 5. METODOLOGÍA</b>	<b>86</b>
5.1 Características de los Participantes	87
5.2 Procedimiento	90
5.3 Calificación de los Exámenes	91
<b>CAPÍTULO 6. RESULTADOS</b>	<b>93</b>
6.1 Comparación dentro Pruebas – Confiabilidad	95
6.2 Comparación entre Pruebas – Validez Concurrente	97
<b>CAPÍTULO 7. CONCLUSIONES</b>	<b>100</b>
<b>BIBLIOGRAFÍA</b>	<b>109</b>
<b>APÉNDICE 1</b>	<b>114</b>
<b>APÉNDICE 2</b>	<b>116</b>
<b>APÉNDICE 3</b>	<b>117</b>
<b>APÉNDICE 4</b>	<b>134</b>
<b>APÉNDICE 5</b>	<b>157</b>

## LISTA DE CUADROS

Cuadro		Página
1	Clasificación ALTE de los niveles de Proficiencia de una segunda lengua	39
2	Formato del First Certificate in English	57

## LISTA DE TABLAS

Tabla		Página
1	Secciones del Examen IELTS	41
2	Relación de puntajes entre algunas pruebas de proficiencia internacionales	46
3	Escala aproximada de puntaje aprobatorio de la prueba FCE	55
4	Reporte sobre confiabilidad consistencia interna del FCE	73
5	Escala de Puntaje de la prueba TOEFL	77
6	Equivalencia de niveles de proficiencia del TOEFL y el CEF	78
7	Formato de TOEFL	79
8	Características de los sujetos participantes	89
9	Puntajes de TOEFL y FCE	94
10	Correlación de puntajes obtenidos en TOEFL y FCE	96

## **1. Introducción**

La necesidad de tener una certificación de conocimientos de una segunda lengua a nivel mundial, se ha convertido en algo común para profesionistas, estudiantes, en los negocios y muchos campos más. La lengua con más demanda de certificación es el inglés y existen dos instituciones a nivel mundial que han destacado en este aspecto, la Universidad de Cambridge en Inglaterra y el Educational Testing Service (ETS) de Estados Unidos. La Universidad de Cambridge cuenta con una amplia variedad de exámenes de proficiencia de la lengua inglesa, de los cuales uno de los más aceptados es el First Certificate in English (FCE). El ETS por su parte tiene también varias pruebas, pero de ellas la más utilizada por universidades es la Test of English as a Foreign Language, más conocida por sus siglas, TOEFL. Los instrumentos de medición que han diseñado, producido y administrado estos dos organismos han sido utilizados para medir la proficiencia de conocimientos de inglés en muchos países. Estas pruebas han sido incorporadas en diferentes tipos de instituciones educativas, y centros de trabajo entre otros, como los instrumentos más comunes para conocer el nivel de proficiencia de candidatos.

### **1.1 Objetivos**

La aceptación de estas pruebas se debe a que se presume que sus componentes tienen validez y confiabilidad. Por esta razón, en el presente

trabajo de pruebas de proficiencia de lengua, se hablará de los diferentes tipos de validez y en particular de la validez concurrente. De esta manera se plantean los objetivos de este trabajo, los cuales son:

1. Conocer la consistencia interna del TOEFL y el FCE por medio de una comparación de estos dos instrumentos de medición.
2. Conocer la validez de estas dos pruebas por medio de una correlación para saber si estas pruebas miden lo mismo, lo que dicen medir y cuál lo mide mejor.
3. Conocer el desempeño de los candidatos en las diferentes secciones de cada una de las pruebas.

Este último objetivo se enfocará en los puntajes obtenidos en cada sección de las pruebas por parte de los estudiantes a quienes se les aplican las pruebas y que son los participantes de la presente investigación. Dichos puntajes se analizarán por medio de estadísticos.

## **1.2 Justificación**

La Universidad de Cambridge y el ETS investigan de manera permanente sus instrumentos de medición para conocer diversos aspectos de los mismos, tales como: niveles de confiabilidad, comparabilidad<sup>1</sup> entre pruebas presentadas en papel y por medio de computadora, análisis de

---

<sup>1</sup> **Comparabilidad.** Posibilidad de equivalencia entre instituciones o entre programas. Implica la exigencia de sistematización, comprensión y claridad.

evaluación de sus pruebas y muchos más.

Pero no se ha documentado una comparación entre las pruebas TOEFL y FCE con respecto a su consistencia interna por medio de una correlación y/o su validez concurrente.

Sólo se tiene información acerca del trabajo que llevaron a cabo Bachman, Davidson, Ryan y Choi (1995) pero éste fue sobre comparabilidad y confiabilidad entre TOEFL y FCE. El objetivo más importante de ese estudio fue la comparabilidad en la medición de las habilidades en ambas pruebas pero también se investigó el desempeño del candidato en las diferentes secciones de TOEFL y FCE, en cuanto a nivel de ítem por ítem. Dentro de sus conclusiones y recomendaciones, estos investigadores dicen que TOEFL tiene un tipo de tareas con un rango poco amplio y que los lineamientos para administrar y calificar la prueba están basados en la estadística y otros aspectos empíricos y cuantitativos. Su sistema de medición sigue la tradición psicométrica, esto permite que TOEFL sea una prueba que minimiza las fuentes potenciales de error. FCE tiene un diseño complejo, cuya administración y calificación son fuente de variaciones en los puntajes de las subpruebas, donde el candidato puede dar más de una respuesta. Los autores recomiendan que deben estudiarse empíricamente estos aspectos del FCE para reducir o controlar estos rasgos de la prueba. Asimismo sugieren adoptar los

lineamientos de las pruebas psicométricas que son más eficientes en limitar la posibilidad de error.

Como se puede ver, el trabajo de los investigadores de referencia ofrece la posibilidad de estudiar diferentes facetas de estas dos pruebas y en diferentes dimensiones, y es así que se decidió llevar a cabo la comparación de ambas para conocer su validez concurrente, su consistencia interna y el desempeño de los candidatos a través de los puntajes obtenidos.

A manera de aclaración, es importante establecer que la presente tesis consiste en la comparación de los puntajes obtenidos en dos pruebas de proficiencia de una segunda lengua, el inglés, y no en el análisis de la validez y la confiabilidad de las mismas, pero por razones de carácter logístico no fue posible modificar el título con el cual se está presentando este documento. No obstante, también es necesario hacer notar que una comparación de este tipo no se ha hecho en México, tampoco han participado estudiantes mexicanos en un estudio a nivel internacional para esta clase de investigación, por lo tanto la información emanada de este ejercicio puede resultar interesante para posteriores trabajos de investigación en nuestro país.

### **1.3 Hipótesis**

Se postulan las siguientes hipótesis para este trabajo.

1. Se corroborará, por medio de los puntajes obtenidos, que la consistencia interna del TOEFL en tres de sus subpruebas, es más fuerte que en cuatro subpruebas del FCE.
2. Los puntajes obtenidos revelarán que la prueba FCE tiene una validez con una alta significancia a diferencia del TOEFL
3. Los candidatos tendrán un mejor desempeño en la prueba FCE al haber tomado preparación para dicho examen.

## **2. FUNDAMENTOS DE PRUEBAS DE PROFICIENCIA DE LENGUA**

En el presente capítulo se expondrán las características más importantes de las pruebas de proficiencia en una segunda lengua más comunes. También se hablará sobre los diferentes tipos y conceptos de validez y confiabilidad en relación a pruebas de proficiencia de lengua principalmente.

## **2.1 Características y Tipos de Pruebas de Proficiencia de Lengua**

Cuando se habla de un procedimiento de medición, diseñado para medir una o diferentes variables de una materia o tema específico, generalmente se habla de una prueba. En cuanto a las pruebas de lengua, éstas coinciden con cualquier otra prueba implementada en una institución en cuanto a los principios generales y hasta en ciertas formas de medición. Sin embargo el medir una lengua no es tarea sencilla debido a que las pruebas de lengua están diseñadas para obtener una muestra específica de un comportamiento que permite hacer inferencias basadas en los resultados de las pruebas. Por lo tanto son instrumentos que permiten reunir evidencia referente a las habilidades generales o específicas con respecto al uso y conocimiento de una lengua por parte de quienes tomen las pruebas. La adecuada interpretación de los datos arrojados por una prueba de lengua contará tanto como evidencia de la proficiencia de quien haya tomado la prueba y como un recurso para obtener muestras que permitan constatar que ésta cumple con los propósitos de medición, aún teniendo como objetivo tan sólo la investigación de habilidades muy generales o globales, según McNamara (2000).

Una instancia de este tipo de pruebas lo constituye el Test of English as a Foreign Language o TOEFL, de acuerdo con Bachman (1990), quien dice que dicho instrumento es considerado un ejemplo de una prueba de

opción múltiple, que proporciona puntajes de pruebas de lengua que son confiables, significativos y útiles.

Existen varios tipos de pruebas de lengua dependiendo del propósito de las mismas, de acuerdo con Alderson, Clapham y Wall, (1996), dichas pruebas se pueden clasificar a grandes rasgos de la siguiente manera:

### **1. Pruebas de colocación**

Este tipo de prueba es utilizado principalmente por instituciones para ubicar el nivel de conocimientos de los candidatos para un curso de lengua.

### **2. Pruebas de avance (Progress tests)**

Son aquellas que se aplican durante el curso para conocer lo que los alumnos han aprendido en las diferentes etapas de estudio del curso.

### **3. Pruebas de logro (Achievement Tests)**

Estas pruebas se aplican al finalizar un curso y se basan en el programa de estudios o libro de texto que se haya llevado

### **4. Pruebas para Propósitos Específicos**

Estas pruebas se basan en un análisis de necesidades para conocer el tipo de lenguaje requerido para un propósito específico como sería el caso del estudio del tipo de lenguaje utilizado por los controladores del tráfico aéreo en los aeropuertos.

## **5. Pruebas de diagnóstico**

Este tipo de pruebas son muy generales pero permiten conocer en que aspectos el candidato necesita reforzar sus conocimientos de la lengua.

## **6. Pruebas de Proficiencia**

Estas pruebas no se basan en un programa de estudios específico ya que están diseñadas para conocer las habilidades de los candidatos con diferentes tipos de preparación y estudio de la lengua. El objetivo es saber si los examinados tienen un cierto nivel general de habilidad en el uso de la lengua para propósitos generales o bien académicos como sería en el caso de medicina, turismo, etc.

En todo tipo de pruebas, y en particular en las pruebas utilizadas como instrumento de medición de conocimientos de una lengua, la tarea de su elaboración requiere de una cuidadosa planeación y consideración de las variantes implícitas para así obtener resultados confiables y válidos.

En este trabajo, cuyo objetivo se centra en las pruebas internacionales de proficiencia de lengua en inglés TOEFL y First Certificate in English, FCE, primeramente se dará una visión general de los aspectos más relevantes en cuanto a validez y de confiabilidad que debe contener una prueba.

## 2.2 Tipos de validez

En situaciones cotidianas, de acuerdo con Anastasi y Urbina (1998), se dice que algo es válido cuando existe evidencia de la veracidad o bien la significación que permite elaborar un juicio al respecto, ya que está bien fundamentado y tiene, o bien una teoría válida, o una razón válida. La validez de los instrumentos de medición tiene que ver con lo que miden y con qué tan bien lo hacen, además de que a través de sus resultados indican qué se puede inferir. Sin embargo los autores recomiendan ser cautos, ya que no se puede aceptar que una prueba, tan solo por su nombre, nos puede indicar lo que mide.

La validez de una prueba se establece, según Cohen y Swerdlik (2000), examinando las fuentes objetivas de información y las operaciones empíricas utilizadas para la elaboración de la prueba para de esta manera establecer su validez. Éste concepto se aplica a una prueba en referencia a un juicio concerniente a lo bien que mide lo que pretende medir. Existe, por lo tanto, un juicio inherente a lo que la prueba en su diseño intenta medir. Se trata, en suma, de un juicio basado en evidencia sobre lo apropiado de las inferencias extraídas de las puntuaciones de la prueba. La validez es, por lo tanto, una cuestión de grado según Messick (1989) y la validación de un test es un proceso continuo, nunca acabado.

La validez es considerada como un juicio evaluativo integrado, que depende del grado en que los fundamentos teóricos y la evidencia

empírica, respalden lo adecuado y apropiado de las inferencias, así como de los análisis de los puntajes de las pruebas. La validez también se enfoca en la identificación de los factores que producen una varianza confiable en los puntajes obtenidos en las pruebas. Socialmente, la validez tiene un valor sobresaliente ya que tiene una función tanto científica como política, que cumple con un coeficiente de correlación entre las puntuaciones de prueba y un pretendido criterio (es decir, la validez clásica relacionada con un criterio) o por juicios expertos de que el contenido de la prueba es relevante para el uso propuesto de la prueba. En otras palabras, se hace referencia a la validez de contenido tradicional, (Anastasi y Urbina 1998). Por otra parte, Aiken y Lewis (1996) encuentran que es una desventaja considerar que la validez sea definida como el grado en el cual una prueba mide algo, ya que la implicación de lo que es una prueba conlleva más de un tipo de validez y ello dependerá de los propósitos específicos con los que se diseñó, la población a la que está dirigida y el método utilizado para determinar la validez de esa prueba. Consideran estos estudiosos que dentro de los métodos para evaluar la validez se encuentran el análisis de contenido, la relación de las calificaciones de la prueba con las calificaciones en base a un criterio y la investigación de los constructos que mide la prueba en cuestión. Este conjunto de consideraciones permite una mejor comprensión de lo que mide una prueba y mejor comprensión de los resultados para formar juicios o tomar decisiones. Otro aspecto a considerar, según estos

autores, es qué tanto aporta una prueba en particular a la predicción y comprensión de los criterios que ya se conocen, para tener así una mejor noción de la validez.

Otro concepto importante de validez es el que ofrece la American Psychological Association, Según esta Asociación la validez debe considerarse como un concepto unitario, aunque argumentan que la evidencia se puede acumular de muchas maneras y que la validez siempre se refiere al grado en que esa evidencia respalda las inferencias que se hacen de los puntajes, es decir que se habla de las inferencias en relación a los usos específicos de una prueba y son las inferencias las que se validan, y no la prueba en sí misma. Por esta razón ven a la validación de una prueba como un proceso de recopilación de evidencia que permite hacer inferencias. Por lo tanto dicen que quienes elaboran una prueba deben suministrar suficiente evidencia de validez en el manual de la misma.

Para tener un concepto más preciso de validez en este proceso de validación, Cohen y Swerdlik (2000) conceptualizan la validez con una perspectiva 'trinitaria' la cual es una taxonomía de tres categorías para clasificar la validez y que ha sido, desde la época de los cincuentas, la perspectiva que prevalece en el campo de la psicología, aún con la inconformidad de algunos investigadores ya que lo consideran un enfoque fragmentado e incompleto. Dicha división de categorías es como sigue.

## **2.3 Validez Aparente**

También llamada validez de Facié no se refiere a lo que la prueba verdaderamente mide sino más bien lo que parece medir. Es un rasgo deseable de las pruebas. En este caso, este tipo de validez indica si la prueba tiene “la apariencia de válida” ante quienes tomarán la prueba, los que la aplicarán o algunos otros involucrados en la aplicación de la prueba. La principal característica de este tipo de validez es el aspecto de las relaciones públicas ya que no bastará que la prueba sea válida para que sea aceptada por los examinados el personal administrativo o bien legalmente, puede haber resistencia ante la prueba por carecer de validez aparente, Anastasi Urbina (1998).

## **2.4 Validez de Contenido**

La validez de contenido tiene que ver tanto con quienes elaboran la prueba como con quienes la toman. Para demostrar que una prueba es relevante y que cubre el área de contenido o habilidad que se supone debe cubrir, requiere ser validada. Existen dos aspectos en este tipo de validez los cuales son: primeramente, la relevancia de contenido y el segundo aspecto el alcance del contenido (Messick,1989). Por lo tanto de acuerdo con estos aspectos Bachman, (1990) haciendo referencia a Popham (1978), propone tener en cuenta los siguientes cuestionamientos a fin de lograr una prueba con validez de contenido

- 1) ¿Qué mide la prueba?

2) Las características de los reactivos que le serán presentados a quien tome la prueba.

3) La naturaleza de la respuesta que se espera obtener por parte de quien tome la prueba.

Un segundo aspecto es propuesto por Bachman y es el referente al alcance del contenido, o sea qué tan bien representa la prueba la cuestión del dominio del comportamiento. En el caso de las pruebas de lengua rara vez se tiene una definición de dominio que de manera clara identifique por medio de un conjunto de tareas el uso real, en su totalidad, de la lengua. En la validez de contenido normalmente encontramos la opinión de los expertos después de analizar el contenido de una prueba y compararla con las especificaciones con respecto a lo que debe contener. Dichas especificaciones pueden estar basadas en la currícula o bien el programa de un curso. Con respecto a los procedimientos para la validación de contenido estos comprenden varios aspectos, principalmente el examen sistemático del contenido de la prueba para determinar si cubre una muestra representativa del área de conducta que debe medirse. Esta forma de validación es utilizada esencialmente para medir qué tan bien ya ha dominado un individuo una habilidad o un curso de estudio. Por ejemplo en el caso de un examen educativo bien formulado éste debe cubrir los objetivos de la instrucción y no sólo su temario. Consecuentemente,

se debe definir el contenido de manera amplia, es decir que éste debe incluir la aplicación de principios y la interpretación de datos. El manual de una prueba educativa de aprovechamiento, por ejemplo, deberá contener la suma de la validación de contenido, además debe incluir la descripción de los procedimientos seguidos para asegurar que el contenido del instrumento es apropiado y representativo. Deberá agregarse información de los expertos consultados y de todos los participantes en la elaboración de la prueba incluyendo información acerca de su calificación profesional (Anastasi y Urbina, 1998).

## **2.5 Validez de Constructo**

Dentro de las múltiples contribuciones de los investigadores, la definición de Messick (1980) es una de las más aceptadas ya que logra reunir las diferentes conceptualizaciones de este tipo de validez. Este autor considera que la validez de constructo es un concepto unificador de validez que integra las consideraciones de contenido y de criterio. Asimismo ofrece una interpretación de los tests basada en la evidencia que proporcionan las puntuaciones y las correlaciones teóricamente relevantes.

Se entiende por constructo una idea científica que se elabora o construye con el fin de describir o explicar un comportamiento. Por lo tanto un constructo es un rasgo presupuesto o bien subyacente, que no

es observable de manera directa. Por ejemplo, cuando se habla de “inteligencia”, se invoca la descripción de por qué un estudiante tiene un buen desempeño en la escuela. Cuando se analiza la validez de constructo de una prueba se debe formular una hipótesis con respecto al comportamiento esperado de quienes obtengan puntuaciones altas y de quienes obtengan puntuaciones bajas. Cuando estos puntajes no corresponden al comportamiento tal y como se había predicho, el investigador deberá volver a examinar las hipótesis planteadas de constructo o bien analizar nuevamente la naturaleza misma del constructo. Esta situación, es muy probable que se deba a que la prueba no sea una medida válida del constructo, o bien la teoría que originó el constructo deba examinarse. Más aún podría existir una ejecución incorrecta del procedimiento estadístico que es lo que proporcionará la oportunidad para encontrar formas alternativas para medir el constructo, (Anastasi y Urbina, 1998). Por lo tanto la validez de constructo mide un constructo o rasgo teórico, como en el caso de la medición de la fluidez verbal en pruebas de proficiencia de lengua.

También Cohen y Swerdlik (2000) consideran que la validez de constructo es el proceso de recolección de evidencia que apoya el hecho de que una prueba realmente mide el constructo psicológico que quienes elaboraron la prueba quieren medir. Otro objetivo de este tipo de validez es determinar el significado que tienen los puntajes obtenidos de la prueba a fin de asegurar que significan lo que se esperaba de ellos, es decir, que

el desempeño de la prueba puede interpretarse como una forma de medir de manera significativa alguna característica o cualidad.

La validez de constructo se ha visto como un concepto unificador de toda la evidencia de validez, como es el caso de validez de contenido y de criterio, a las cuales se les puede considerar como validez de constructo también.

Existen otros dos tipos de validez, la convergente y la discriminante, las cuales Campbell y Fiske (1959) consideran que forman parte esencial del proceso de validez de constructo ya que argumentan que una prueba, desde su diseño, debe tener variables que puedan correlacionarse o bien discriminarse.

Los autores definen a estos tipos de validez de la siguiente manera.

### **2.5.1 Validez Convergente**

Cuando existe una prueba que tiene una correlación elevada con otras variables que teóricamente deben correlacionarse, existe una validez convergente. Un ejemplo de este tipo de validez lo proporcionan Anastasi y Urbina (1998) cuando dicen que si se comparan los puntajes obtenidos en una prueba de razonamiento cuantitativo con las calificaciones logradas en una prueba de matemáticas se trata de una validez convergente. En este ejemplo se trata de pruebas que contienen variables relevantes y que se espera que deban converger al ser comparadas.

### **2.5.2 Validez Discriminante**

En este caso no existe una alta correlación con variables que en teoría deben diferir, como por ejemplo en el caso de existir una correlación baja o no significativa entre una prueba de razonamiento cuantitativo y una prueba de comprensión de lectura. La razón por la que se espera que no haya una correlación significativa y que las variables difieran en este ejemplo, es porque la variable referente a la habilidad de lectura es irrelevante en una prueba que mide el razonamiento cuantitativo, (Anastasi y Urbina, 1998).

## **2.6 Validez Relativa al Criterio**

El enfoque de este trabajo es sobre validez, pero particularmente la validez en su aspecto concurrente, la cual es una subdivisión de la validez relativa al criterio.

Existen dos características importantes para la elaboración de pruebas, psicológicas y educativas, y estas son la norma y el criterio. Aiken y Lewis (1996) consideran que una norma es un marco de referencia que se utiliza para interpretar calificaciones brutas. Para obtener dicha norma es necesaria la estandarización en la aplicación y calificación de pruebas a fin de tener una distribución de las calificaciones brutas en un grupo de norma o estandarización. La razón es que los individuos de ese grupo norma que tomen las pruebas en cuestión,

representan a la población meta y las calificaciones brutas obtenidas por ese grupo se convierten en calificaciones derivadas o normas.

A un criterio lo definen los autores como un estándar o variable con el cual se puede comparar el desempeño de una prueba. Es por lo tanto una norma contra la que se valora una prueba o bien la puntuación de una prueba.

El criterio debe ser confiable, relevante y válido ya que puede contaminarse cuando los resultados de la prueba influyen en la condición del criterio de cualquier individuo. Por ejemplo, cuando se otorgan calificaciones laborales o académicas, si un sujeto tuvo un desempeño pobre, o muy bueno, en una prueba determinada y dicho dato es conocido por quien finalmente asignará una calificación, ese conocimiento previo a su juicio puede influir en la calificación que asigne existiendo así ya un criterio contaminado. Por esta razón quienes tengan a su cargo la asignación de calificaciones del criterio no deben tener conocimiento previo de los resultados que los examinados hayan obtenidos en la prueba aplicada, Salvia et al (1997).

Para validar una prueba, ésta puede compararse con cuantos criterios y usos específicos se tengan para dicha prueba. Un ejemplo de este tipo de validación son las pruebas de ingreso a la universidad, donde existe un criterio común el cual se basa en la calificación promedio

obtenida en todos los cursos en el transcurso del primer año de estudios. Cada calificación es calculada por el número de puntos de cada curso y por quienes tomaron cada curso.

La validez Relativa al Criterio tiene dos subdivisiones: validez predictiva y validez concurrente.

### **2.6.1 Validez Predictiva**

Esta validez se refiere al grado en que las puntuaciones del test pueden predecir medidas de criterio tomadas posteriormente, es decir, las conductas futuras del sujeto en el criterio, como por ejemplo en el caso de las pruebas de admisión en una institución educativa, las puntuaciones del test inicial pueden proyectar cuales serán a futuro los resultados en un test de desempeño académico. Otro uso frecuente de este tipo de validez se da en el campo de la medicina para tratamientos médicos, en la industria para promoción de empleados, etc. (Anastasi y Urbina, 1998)

En el caso de pruebas de proficiencia de la lengua para fines de académicos o laborales, este tipo de prueba es muy común ya que permite tener una predicción con respecto al desempeño futuro del sujeto en un ambiente donde tendrá que demostrar sus habilidades en el uso de la lengua.

## 2.6.2 Validez Concurrente

Primeramente se debe decir que la validez convergente y la concurrente tienen en común que se analizan por medio de las correlaciones de los puntajes que se obtienen de las pruebas que se estén estudiando. Existe validez convergente cuando los puntajes de una prueba tienen una correlación alta con otros puntajes de pruebas que evalúan el mismo tipo de constructos. La validez concurrente involucra la comparación de los puntajes obtenidos en una prueba con otra medida similar o paralela, pero a diferencia de la validez convergente se trata de dos pruebas aplicadas a los mismos sujetos y aproximadamente en el mismo período de tiempo de acuerdo con Bachman (1990).

Otra característica que este autor agrega respecto a la validez concurrente es que los puntajes deben expresarse de manera numérica ya que el resultado de la comparación generalmente se expresa por medio de un coeficiente de correlación. En pruebas de lengua este tipo de validez es muy común y la información obtenida se clasifica de dos maneras:

1. “Se examinan las diferencias en el desempeño en una prueba entre grupos de sujetos con diferentes habilidades en el manejo de la lengua.”
2. “Se examinan las correlaciones entre varias mediciones de una habilidad determinada. Si se puede identificar a grupos de

sujetos que tengan diferentes niveles en la habilidad que interese medir, se puede entonces investigar el grado en el que el test de esta habilidad discrimina de manera precisa a los sujetos de los grupos participantes. Regularmente participan en este tipo de pruebas tanto nativo hablantes como no nativo hablantes de la lengua que se trate en la prueba.”

Cuando se produce un nuevo test frecuentemente se busca saber qué correlación tiene en cuanto al puntaje con una prueba ya existente y considerada como estandarizada. Pero una seria limitación en el caso de la validez concurrente lo constituye el hecho de que no se esperaría tener una correlación de puntaje entre una prueba de matemáticas y una prueba de fechas importantes en la historia, así que para demostrar que los puntajes obtenidos en una prueba son indicadores válidos de una habilidad en una lengua se debe demostrar que se relacionan con otros indicadores de la misma habilidad y que no están relacionados con la medición de otras habilidades. Se puede mostrar también la existencia de validez concurrente correlacionando los puntajes obtenidos por los estudiantes con los puntajes asignados por sus maestros de lengua, siguiendo para ello un formato específico para calificar las diferentes habilidades de los alumnos, (Bachman, 1990)

Por todo lo anterior, en este trabajo se pretende conocer la validez de dos pruebas de proficiencia de lengua inglesa, TOEFL y FCE, las

cuales se supone contienen variables cuyos puntajes se pueden correlacionar y así conocer a su vez, como se correlacionan ambos instrumentos.

## **2.7 Confiabilidad**

Existen dos aspectos fundamentales a considerar en una prueba y ellos son confiabilidad y validez. Cuando se habla de confiabilidad se refiere a una medida consistente de medición que permite obtener resultados similares de manera, a su vez consistente, en diferentes contextos. Sin embargo la confiabilidad también está sujeta a factores que nada tienen que ver con las habilidades como sería el caso de instrucciones confusas contenidas en la prueba, ruido en el área de exámenes o bien poca claridad en la calificación e interpretación de los resultados, entre otras, pero que afectan el desempeño y el puntaje de quienes toman una prueba de conocimiento de una lengua. Al minimizar este tipo de variables se podrán tener resultados más confiables. Otro factor es la heterogeneidad de quienes tomen la prueba y la cantidad y dificultad de reactivos contenidos en la prueba por lo que para fines confiables es deseable tener una dificultad media (Bachman y Palmer, 1996).

No obstante, la confiabilidad no asegura que exista validez en una prueba, es decir el grado en que una prueba mide aquello para lo que se diseñó. Esto se puede ejemplificar de la siguiente manera: Supongamos

que a una báscula digital se le considera confiable ya que mide de manera consistente en la misma forma, cuando se hacen mediciones repetidas de un mismo peso estándar, como una barra de oro de un kilogramo por ejemplo obteniendo siempre el mismo resultado, razón por la que se diría que se trata de un instrumento de medición confiable. Si la misma barra de oro se pesa en otra báscula digital y se obtiene de manera repetida un resultado diferente, 1.3 kilogramos, se diría que el resultado es impreciso e inválido aunque confiable. Por lo tanto confiabilidad en un sentido psicométrico únicamente denota algo consistente, no necesariamente bueno o malo, tan sólo consistente según Cohen, Swerdlik y Mark, (1996).

Existen diferentes tipos y grados de confiabilidad pero en este trabajo solo se hará referencia a algunas de las más comunes en las pruebas de lengua.

## **2.8 Confiabilidad Interjuez o entre evaluadores**

Este tipo de confiabilidad es muy utilizada en las pruebas de proficiencia de lengua inglesa que elabora la Universidad de Cambridge, en los exámenes de producción escrita y de expresión oral. La confiabilidad interjuez es el grado de acuerdo que existe entre jueces o calificadores para una prueba en particular. Ya sea por medio del manual o algún tipo de instructivo de la prueba, los calificadores son capacitados para que su desempeño sea altamente homogéneo y sus juicios se basen

en los parámetros de calificación especificados en el manual de la prueba. Para corroborar el grado de consistencia entre evaluadores en referencia a la calificación de una prueba se calcula un coeficiente de correlación por medio de instrumentos estadísticos, debiendo así dos de los jueces coincidir en su evaluación.

## **2.9 Estimaciones de Confiabilidad de Prueba y Posprueba**

Para estimar la confiabilidad de un instrumento de medición se debe usar ese mismo instrumento para medir lo mismo en dos puntos en el tiempo. A esto se le llama Estimaciones de Confiabilidad de Prueba y Posprueba. En el caso de pruebas de lengua la aplicación de una prueba por segunda ocasión y con los mismos sujetos, puede arrojar resultados similares. Pero existe una limitante, ya que aún con un breve intervalo entre una y otra prueba puede haber baja confiabilidad ya sea porque las personas resienten volver hacer la prueba o bien recuerdan los ítems y ya tienen respuestas diferentes o más aún las personas cambian ya que con el tiempo reciben más información, amplían sus conocimientos, etc., por lo que su desempeño se ve afectado por esas circunstancias, Anastasi y Urbina (1998).

## **2.10 Confiabilidad: Formatos Paralelos**

Otra forma de establecer la confiabilidad de una prueba es la de formatos paralelos, la cual implica la correlación de puntajes en pruebas

muy similares o paralelas pero realmente el hecho de elaborar dos pruebas con características tan similares que se les puedan clasificar como paralelas es una tarea sumamente difícil, casi imposible. Dentro de las características importantes de confiabilidad de una prueba de este tipo, se pueden mencionar la extensión de la prueba y el rango de habilidades de los alumnos que tomarán la prueba. También debe considerarse la homogeneidad de los ítems los cuales influyen de manera directa en la confiabilidad, como por ejemplo, cuando todos los ítems se elaboran para probar la misma habilidad de la misma manera, entonces los ítems tienen una alta intercorrelación lo que origina un alto índice de confiabilidad, (Alderson, Clapham y Wall, 1995). Los autores de referencia también hacen notar que se debe recordar que la confiabilidad no se mide sino más bien se estima y para ello también existe la confiabilidad en relación a la consistencia interna, la cual será utilizada en este trabajo cuando se correlacionen las pruebas TOEFL y FCE.

### **2.11 Confiabilidad: Consistencia Interna**

Este tipo de confiabilidad lo define Bachman (1990) como la consistencia en el desempeño de los candidatos con respecto a las diferentes partes que constituyen el instrumento de medición. El grado en que una prueba es confiable puede estimarse por medio del coeficiente de confiabilidad,  $r$ , es decir en términos de porcentajes, como explica Nunan (1995), dicho coeficiente funciona como un coeficiente de correlación el

cual puede llegar hasta +1.0 en el caso de una prueba completamente confiable. Este mismo autor expone que el coeficiente de confiabilidad estima el porcentaje de la variación en los puntajes obtenidos en referencia al criterio, añade que la estimación de la consistencia interna de la prueba dependerá del porcentaje de la variación confiable en los puntajes. La razón de esta estimación se debe a que en las ciencias sociales se considera que si se obtiene un  $r$  de 0.05 es significativo pero cuando existe un nivel de 0.01 es altamente significativo y por lo tanto se podría estimar que la consistencia interna de una prueba con esos valores, es alta. Por lo tanto en la comparación que se hará de las pruebas TOEFL y FCE, se respetarán estos criterios para conocer la confiabilidad con respecto a la consistencia interna de estas pruebas.

Es importante agregar que en el caso de las pruebas de lengua, existe una ambigüedad entre confiabilidad y validez ya que la distinción entre uno y otro conceptos no se puede aplicar en todos los contextos debido a que la interpretación que se le da a los puntajes obtenidos en las pruebas depende de los usos para los cuales se diseñó la prueba; por ejemplo, cuando una prueba es calificada principalmente en cuanto a que el contenido sea correcto y completamente contestado, como en el caso de una prueba de escritura, se podría encontrar que la medición no es confiable en cuanto a la habilidad para la escritura del candidato pero si

puede ser válida en cuanto a que el individuo si tiene los conocimientos del tema especificado en la prueba, Bachman (1990).

Un ejemplo de este tipo de problemática son las pruebas de escritura que forman parte de las pruebas de proficiencia de segunda lengua y en las cuales los candidatos plasman sus conocimientos tanto lingüísticos como en referencia a los temas acerca de los cuales deben redactar documentos. La definición de qué se está evaluando en este tipo de pruebas incluye aspectos que involucran el conocimiento de la lengua meta, el contenido de los textos redactados por los candidatos, el conocimiento de las convenciones de redacción de documentos en la lengua meta, entre otros aspectos que en su conjunto complican establecer con claridad si se trata de confiabilidad o validez.

### **3. PRUEBAS DE PROFICIENCIA DE CONOCIMIENTOS DE LENGUA**

En este capítulo se hablará de algunas pruebas de proficiencia de lengua que forman parte de los requisitos para ingresar a una institución académica, obtener una beca, un empleo, entre otros propósitos. Asimismo se expondrá cómo se conceptualiza el aspecto de la proficiencia en una lengua y lo que implica, las diferentes propuestas por parte de los investigadores que a pesar de los trabajos realizados en las últimas décadas, aún no han podido llegar a un consenso debido a la complejidad de este constructo.

### **3.1 Definición de proficiencia de lengua**

Dos investigadores, Oller y Perkins (1980), comentan que durante mucho tiempo se pensó que la proficiencia en una lengua consistía en un constructo global único que era unidimensional y que no comprendía elementos pragmáticos, psicolingüísticos o socioculturales, pero desde los años 80s y debido a la necesidad de medir los conocimientos de segundas lenguas, los investigadores y quienes elaboran este tipo de pruebas se han cuestionado cómo medir y calificar todo lo que involucra conocer y usar una lengua en diferentes ámbitos y circunstancias. Por lo tanto como resultado de diversas investigaciones se encontraron evidencias que permitieron sugerir que se trataba de un constructo multidimensional tal y como los investigadores de referencia concluyeron después de extensos análisis de diferentes pruebas de proficiencia de lengua. Otra propuesta para definir a la proficiencia es la de Cummins (1980) quien propuso adoptar dos conceptos de proficiencia y para tal fin sugiere el análisis de las tareas implicadas en una prueba de proficiencia a través de dos dimensiones: “El rango del soporte contextual” y “El grado de la implicación cognitiva”. Esto resultaría en una proficiencia llamada “Habilidades comunicativas interpersonales básicas” que según este autor son habilidades que se adquieren fácilmente, y un segundo concepto sería la “Proficiencia cognitiva/académica de la lengua” la cual es más difícil de adquirir. Es decir que el autor define a la proficiencia, sumando los dos

conceptos, como el resultado de la conjugación de diversos elementos de la lengua que permiten la comunicación efectiva en un contexto específico.

Hoy en día la producción de pruebas para determinar el nivel de conocimientos que tanto inmigrantes como estudiantes extranjeros tienen de una lengua, así como las pruebas internacionales de proficiencia de una lengua, elaboradas en la Unión Europea, y América del Norte, no se basan en programas de estudios específicos sino que tratan de medir básicamente lo que el individuo puede hacer con el idioma. La proficiencia en una lengua para este objetivo, se vislumbra como la suma de habilidades que el individuo ha desarrollado para usar una lengua y que fundamenta el proyecto “can do” el cual ha impactado prácticamente en todo el mundo, desde la metodología de la enseñanza de segundas lenguas, producción de textos, exámenes etc.

Este concepto ha sido utilizado por el Marco Común de Referencia Europeo, cuyas siglas son CEF por su nombre en inglés, Common European Framework. Dicha institución forma parte de la División del Lenguaje del Consejo Europeo el cual busca facilitar la movilidad internacional europea, satisfaciendo al mismo tiempo las necesidades multilingües y multiculturales de los europeos a través del desarrollo de las capacidades comunicativas de los hablantes, a fin de diversificar e intensificar el aprendizaje de lenguas para promover el plurilingüismo (Council of Europe, 2003<sup>a</sup>). Esta política ha llevado a la clasificación de

los niveles de conocimientos de una segunda lengua en base a estándares que describen lo que los sujetos pueden hacer con el idioma en las cuatro habilidades: escritura, lectura, habilidad oral y conocimientos de gramática (Clark, 2002). Pero a pesar de la aceptación de este tipo de medición de conocimientos de una lengua, algunos investigadores como Baker (1997) argumentan que la lengua no consiste sólo en habilidades sino que involucra cuestiones cualitativas como por ejemplo las funciones poéticas y emotivas que existen en las diferentes lenguas, las cuales no pueden medirse fácilmente en una prueba.

Para fines prácticos las instituciones educativas y las instancias de gobierno en países con un nivel de inmigración alto, como los Estados Unidos y Canadá, que requieren tener parámetros que les permitan conocer que tan proficiente en la lengua meta es un no nativo hablante, han tenido que elaborar una definición de lo que es un individuo proficiente en el uso de la lengua inglesa. Por ejemplo la Legislación Federal de Educación de los Estados Unidos y los académicos con fundamento en la Ley para la Educación Bilingüe, han considerado que existen dos tipos de proficiencia, desde el punto de vista académico: cuando la persona es proficiente de manera limitada y el otro cuando la persona tiene una fluidez proficiente en inglés. Al proficiente limitado lo definen de manera general, Del Vecchio y Guerrero (1995) de la siguiente manera: Este tipo de hablante es aquel individuo que aún siendo nativo de los Estados

Unidos de América, Alaska o habiendo nacido en un país donde la lengua dominante no sea el inglés, presenta de igual manera dificultad para comprender, escribir y leer la lengua inglesa. Esta circunstancia no le permitirá tener un buen desarrollo académico cuando ingrese a una institución educativa, donde la lengua de instrucción sea el inglés, por lo tanto este individuo no podrá participar ni desarrollarse de manera integral en la sociedad.

Esta definición es compartida por las autoridades educativas correspondientes y conlleva un impacto social, particularmente entre los inmigrantes, ya que esta limitante no les permite tener acceso a la educación monolingüe, por lo tanto las instancias gubernamentales correspondientes han tenido que integrar escuelas con programas bilingües situación que ha obligado a las autoridades del Consejo de Educación de los diferentes estados de la Unión Americana a desarrollar su propia definición de proficiencia en la lengua, la cual se basa en las cuatros habilidades:

- a) **Lectura** – puede comprender e interpretar textos de acuerdo a su edad y nivel de educación.
- b) **Habilidad auditiva** – puede comprender el lenguaje que utiliza el profesor, instrucciones, puede extraer información y seguir el discurso en el aula.

- c) **Escritura** – puede producir textos con un contenido formal lo que le permite elaborar trabajos de acuerdo a su edad y nivel de educación.
- d) **Expresión Oral** – puede utilizar el lenguaje oral de manera apropiada y efectiva en las actividades de aprendizaje tanto dentro del salón de clases como en la interacción social dentro de la escuela.

Con respecto a los individuos que son considerados como proficientes limitados, tienen las mismas características que las citadas por la Legislación Federal de Educación. Es decir que la proficiencia en una lengua es vista como todo un conjunto de elementos que implican desde vocabulario, gestos y cuestiones discursivas que permiten la comunicación de manera significativa en un contexto específico, en este caso el contexto es precisamente la escuela.

Dentro del concepto de proficiencia Butler y Hakuta (2004) incluyen las cuestiones del conocimiento pragmático y la habilidad para usar dicho conocimiento, así como el desempeño dentro de los contextos socio-culturales propios de los nativos hablantes de la lengua meta, aspectos que marcan una característica más que considerar dentro del concepto de proficiencia ya que la competencia comunicativa no se limita a conocer los códigos lingüísticos de los significados referenciales. A este respecto,

Shrum y Glisan (2005) retoman la investigación de Magnan (1986) quien habla del proceso de adquisición de una lengua en referencia a la proficiencia, y reporta que encontró que existen diferentes niveles de proficiencia que se entrelazan y forman un continuo. Dichos niveles dependen de la etapa de aprendizaje-adquisición en la que el sujeto se encuentre y no de un programa de estudios o un límite de tiempo establecido para el estudio de una lengua. También Cummins (2000) propone que la proficiencia se desarrolla en un continuo que se prolonga más allá de las aulas y continúa durante la vida, ya que el desarrollo del lenguaje se caracteriza por el incremento en la diferenciación que el individuo puede hacer al utilizar el lenguaje de acuerdo a las tareas a realizar en los diferentes contextos. Dicha diferenciación es producto del conocimiento que tiene el individuo de los registros lingüísticos, la fonología, etc., elementos específicos que le permiten funcionar de acuerdo al contexto. Pero como la proficiencia en este concepto va de la mano de la instrucción escolarizada propia de las sociedades industrializadas donde la educación es obligatoria hasta los 16 años de edad para toda la población, los individuos llegan de esta manera a tener un nivel de proficiencia que se refleja en el grado de dificultad de las tareas que desarrollan tanto dentro del campo educativo como fuera de él. Cummins (2000) comenta que Oller (1979) interpreta este aspecto como una Proficiencia General de la Lengua y califica a este tipo de proficiencia como el constructo de la Proficiencia Cognitiva/ Académica del Lenguaje

(*CALP*). Esta última toma más tiempo en ser desarrollada por los hablantes y ese desarrollo depende mucho de los contextos donde se adquiera ya que según Cummins no existe una estructura de la proficiencia que se pudiera considerar universal y tampoco fuera de un contexto.

Sin embargo Cummins sugiere diferenciar la proficiencia de la lengua desde el punto de vista académico *CALP* y la Proficiencia Académica/ Conversacional de la Lengua. Este último aspecto demanda el uso de las características cognitivas dentro del contexto donde se desarrolle el discurso. Este constructo psico-educacional no es inferior ante *CALP*, lo que ha permitido hacer una distinción de estos dos tipos de proficiencia, que son:

1. En los contextos monolingües esta distinción se refleja en la proficiencia de la lengua adquirida por medio de la interacción como sería el caso de un niño de seis años y la proficiencia adquirida a través de la instrucción escolarizada la cual continúa por toda la vida de los hablantes.
2. En estudios hechos desde principios de los 80s, se ha demostrado que los estudiantes inmigrantes al estar expuestos al contexto del salón de clases y a diversos contextos dentro de su medio ambiente, rápidamente logran una fluidez en la lengua

meta, pero les toma hasta más de cinco años lograr una proficiencia en cuanto al aspecto académico con respecto a los nativos hablantes de la lengua meta.

Con respecto a los bilingües, la propuesta de Cummins quiere dejar en claro que la proficiencia de la lengua no quiere decir únicamente que el sujeto es capaz de leer y escribir y que tiene conocimientos básicos de la lengua meta. Agrega que el conocimiento de la lengua en cuestión se manifiesta en su habilidad conversacional y de escritura para expresarse en contextos académicos, ya que según este autor existe una discriminación potencial hacia los bilingües inmigrantes cuando se toma su fluidez conversacional en la lengua meta, como índice de su adquisición académica en dicha lengua.

Como se puede ver la cuestión de la proficiencia en una lengua puede tener repercusiones importantes para los candidatos interesados en estudiar o emplearse. En Europa han respondido a esta problemática organizándose y trabajando conjuntamente y para ello fundaron la Asociación de Examinadores de Lenguas de Europa, *ALTE*, la cual es una asociación que cuenta actualmente con 28 países miembros que representan 24 lenguas europeas. Sus parámetros se basan en los principios y objetivos de la CEF y desde su fundación emprendida en 1989 por la Universidad de Cambridge y la Universidad de Salamanca, los

miembros de la asociación han trabajado para fijar los siguientes objetivos en la definición de un marco de referencia común, con respecto a los conocimientos de una segunda lengua que debe tener un candidato, para acceder a una institución de estudios superiores, un empleo, etc., (ALTE, 2004).

1. Establecer niveles de proficiencia de conocimientos de lengua comunes para promover el reconocimiento y certificación en toda Europa.
2. Establecer estándares comunes para todos los niveles del proceso de la elaboración y aplicación de pruebas, diseño de ejercicios y de ítems, evaluación, reporte de resultados, análisis y reporte de hallazgos.
3. Colaborar de manera conjunta en proyectos y en el intercambio de ideas y metodología.

El marco del ALTE (2004) propone seis niveles comunes de referencia para clasificar la proficiencia en una segunda lengua, y cada uno de ellos detalla específicamente lo que puede hacer el usuario de la lengua desde el punto de vista receptivo, de producción y en una interacción. (Véase cuadro No. 1) Dichos niveles van desde la

proficiencia identificada como básica hasta un usuario proficiente como a continuación se explica brevemente, (The Council of Europe, 2003a).

<b>Cuadro No. 1 CLASIFICACIÓN ALTE DE NIVELES DE PROFICIENCIA DE UNA SEGUNDA LENGUA</b>		
<b>Usuario Básico:</b>	<b>A1</b>	<b>Breakthrough o Pre-Waystage</b>
Puede entender y usar expresiones cotidianas familiares y frases muy utilizando lenguaje muy básico, es decir con frases cortas o poco vocabulario. También puede hacer y responder preguntas acerca de detalles personales.		
<b>Usuario Básico:</b>	<b>A2</b>	<b>Waystage</b>
Puede entender expresiones de uso frecuente referentes a información personal relevante. Puede comunicarse en tareas sencillas y de rutina que requieren de un intercambio sencillo, tiene una fuerte influencia de su L1, lo cual puede impedir la comprensión en su conversación. Puede escribir mensajes y correos electrónicos muy cortos.		
<b>Usuario Básico e Independiente</b>	<b>B1</b>	<b>Threshold</b>
Puede entender los principales puntos de una conversación referente a temas como la familia o el trabajo, puede expresar sus puntos de vista aunque con limitaciones cuando requiere utilizar formas complejas del lenguaje en cuanto a estructuras gramaticales o vocabulario. Puede redactar cartas para solicitar empleo, dar información, describir su entorno, etc.		
<b>Usuario Intermedio Independiente:</b>	<b>B2</b>	<b>Vantage</b>
Puede manejar la comunicación de manera eficiente en casi todo tipo de situaciones ya que en ocasiones sus emisiones con incoherentes o tienen errores gramaticales o bien de vocabulario. Puede redactar narraciones, hacer resúmenes, solicitar información, empleo, etc., por medio de una carta,		
<b>Usuario Proficiente:</b>	<b>C1</b>	<b>Efficiency</b>
Puede entender un amplio rango de textos, extensos y sobre diversos temas. Puede usar el lenguaje de manera flexible, efectiva y de ser necesario en un campo profesional. Puede escribir textos extensos mostrando un buen uso de los patrones de organización del discurso.		
<b>Usuario Proficiente:</b>	<b>C2</b>	<b>Mastery</b>
Puede entender con facilidad casi todo lo que escucha y lee. Puede resumir información, puede expresarse de manera espontánea tanto con mucha fluidez como con precisión. Puede reconstruir información de manera coherente diferenciando significados un tanto oscuros en textos y el discurso y en situaciones más complejas que las de los anteriores niveles.		

A continuación se hablará de las pruebas de proficiencia de lengua como L2 y que son de carácter oficial en instituciones universitarias y a nivel internacional, como la prueba TOEFL, por lo que se describirán algunas de las pruebas de proficiencia más utilizadas para determinar el ingreso a una universidad, obtener una beca o bien un empleo.

### **3.2 DIFERENTES TIPOS DE PRUEBAS DE CONOCIMIENTOS DE LENGUA**

En el mundo globalizado de hoy existe la demanda de instrumentos de medición de conocimientos de una segunda lengua, no solo debido a la migración sino también a que las universidades y empresas que ofertan empleos requieren que los candidatos sepan dos idiomas, particularmente cuando la L2 es la utilizada como medio de enseñanza o de comunicación cotidiana, esta circunstancia ha originado el diseño y desarrollo de pruebas que permiten conocer el nivel de proficiencia que tiene el candidato.

Las pruebas de proficiencia de lengua se clasifican por lo general como Pruebas Estandarizadas ya que esencialmente comparan el puntaje obtenido por un sujeto contra el puntaje obtenido por otros, como por ejemplo en las pruebas de inteligencia (Baker,1997). En este trabajo el enfoque es en pruebas de inglés utilizadas a nivel internacional. Una de ellas, la cual es ampliamente utilizada en el mundo para ingresar a una universidad es la llamada Sistema Internacional de Pruebas de Inglés, (International English Language Testing System, IELTS). Esta prueba fue en sus inicios utilizada por el Consejo Británico para determinar si los candidatos tenían un manejo adecuado del inglés para poder acceder a una beca de estudios en el Reino Unido, pero a partir de 1995 este instrumento ha sido modificado y monitoreado constantemente.

El examen tiene un programa permanente para mejorar cualquier aspecto que por medio del departamento de investigación, se recomiende implementar en las nuevas pruebas. Actualmente, al IELTS lo administran y manejan directamente el Consejo Británico, Cambridge ESOL Examinations y el Departamento Internacional de Desarrollo de Programas de las Universidades Australianas de Educación, (DIP). Esta prueba consiste en cuatro módulos, (véase tabla no.1) dos de ellos prueban la comprensión auditiva y de expresión oral, dichos módulos son los mismos para todos los candidatos. Con respecto a los otros dos módulos, comprensión de lectura y escritura, el candidato puede elegir entre módulos que prueben aspectos académicos o una capacitación en general en estas dos habilidades. Las pruebas de este examen han sido diseñadas para que cubran un amplio espectro de habilidades que van desde el clasificado como un usuario inexperto hasta uno experto. El diseño y elaboración de la prueba IELTS se basa en los parámetros especificados por la CEF y ALTE, por lo que su evaluación se reporta en las bandas de clasificación de proficiencia utilizadas por estas organizaciones. La vigencia internacional de esta certificación es de hasta dos años únicamente.

<b>Tabla No. 1 SECCIONES DEL EXAMEN IELTS</b>	
<b>Examen</b>	<b>Duración</b>
Comprensión Auditiva	30 minutos
Comprensión de Lectura	60 minutos
Expresión Escrita	60 minutos
Expresión Oral	Entre 11 y 14 minutos

Duración total del examen: 2 horas 45 minutos

El IELTS contiene 40 reactivos tanto en la sección de Comprensión Auditiva como en la de Comprensión de Lectura y la banda de calificación para estas dos secciones va del 1 al 9. La sección de Expresión Escrita se califica considerando una banda de puntaje con igual peso en cada uno de estos aspectos: cohesión, coherencia, estructuras gramaticales, vocabulario y registro en la redacción de los dos documentos que debe elaborar el candidato. Por último el examen de Expresión Oral también es calificada por medio de bandas de puntaje con igual peso en cada uno de estos aspectos: fluidez, coherencia, vocabulario, estructuras gramaticales, pronunciación y corrección en su expresión oral. Los examinadores de esta última sección reciben una intensa capacitación para asegurar que sus evaluaciones sean validas y confiables.

Específicamente los módulos de expresión escrita y lectura académicos evalúan si el candidato tiene los conocimientos del idioma para ser utilizado en un ambiente formal, académico que le permitan estudiar o trabajar en un contexto donde el inglés sea el medio de enseñanza o comunicación a un nivel universitario, de postgrado y profesional o bien con el propósito de emigrar a Australia, Canadá y Nueva Zelanda. La calificación de aceptación del candidato depende de la institución y objetivo para tomar este examen, en cuanto a una universidad se requeriría regularmente un nivel 6 para licenciatura.

El sistema de calificación de IELTS es revisado y mejorado constantemente y para ello se busca que los puntajes reflejen lo más justamente posible la calificación que obtuvo un candidato. Actualmente todas las secciones se reportan con medios puntos o medias bandas en todas sus secciones con el fin de que el candidato no pierda puntuación por una falta de aproximación en la calificación final de una sección. Por ejemplo, si un candidato obtiene 5.25 en una sección, se le reporta 5.5 o si obtuvo 5.75 se le reporta 6. La organización de IELTS argumenta que este proceso se lleva a cabo por medio de un sistema más específico y fino en cuanto a la calificación de los puntajes (IELTS, 2007).

Este examen es ampliamente aceptado en todo el mundo, razón por la que rivaliza con TOEFL como el examen de proficiencia de lengua requerido por un buen número de universidades americanas, canadienses y europeas. El IELTS también ha implementado el sistema de examen por computadora a partir del año 2005, lo que lo pone a la par de otros exámenes de proficiencia.

En el caso de Canadá, debido al alto índice de inmigración el departamento de Ciudadanía e Inmigración de Canadá (CIC), ha desarrollado sus propios parámetros llamados Parámetros Canadienses de Evaluación de la Lengua (CLBA), con el fin de tener estándares que les permitan medir y describir la proficiencia que en el aspecto de comprensión auditiva, expresión oral, comprensión de lectura y

escritura debe tener un inmigrante en el uso de la lengua inglesa. Estas habilidades se dividen en tres niveles de proficiencia y cada nivel está constituido por cuatro estándares sumando un total de 12 puntos. Estas clasificaciones permiten a los inmigrantes tener acceso a instituciones educativas o bien obtener un empleo, motivo por el cual este sistema de evaluación de proficiencia en la lengua es crucial para los inmigrantes. El CLBA es un sistema que cae dentro del tipo de evaluación comunicativa, tomando en consideración que esta noción se basa en que la lengua es interactiva, directa y auténtica lo cual implica que las pruebas contengan lenguaje que represente una réplica, hasta donde sea posible, de una interacción real utilizando para ello textos auténticos para las pruebas escritas y de expresión oral (Cummins, 2000).

El examen llamado Evaluación de Inglés Académico Canadiense (CAEL), es comúnmente utilizado en Canadá para conocer el nivel de proficiencia del inglés que tiene un estudiante, no nativo hablante, para poder asistir a una universidad o una institución de nivel superior. El objetivo del examen es identificar si el candidato podría tener éxito en sus estudios universitarios desde el punto de vista del buen manejo de la lengua hablada, escrita y leída en un contexto de estudios universitarios. A diferencia del CLBA, este examen es de carácter meramente académico, no para fines de inmigración o empleo. Tampoco los exámenes TOEFL o IELTS tienen similitud alguna con este examen ya

que el CAEL es un examen de proficiencia de lengua, de los clasificados para evaluar el idioma Inglés con Propósitos Académicos (EAP), y su diseño está basado en temas académicos a nivel universitario como ya se mencionó anteriormente, asimismo su diseño es diferente al de otras pruebas como el TOEFL, por ejemplo, ya que no hay tareas de opción múltiple.

El CAEL fue desarrollado en 1987 por la Universidad de Carleton y desde 1989, año en que se usó por primera vez, se ha aplicado como requisito de inscripción a estudiantes que desean cursar estudios universitarios principalmente en Canadá, Estados Unidos de Norteamérica y Europa. Su manejo y administración está a cargo de la School of Linguistics and Applied Language Studies de la Universidad de Carleton en Canadá.

La prueba tiene una duración de 2 horas con 20 minutos y está dividida en cuatro secciones:

- Comprensión de Lectura: con una duración de 50 minutos
- Comprensión Auditiva: con una duración de 20 minutos
- Escritura: con una duración de 45 minutos
- Expresión oral: con una duración de 25 minutos

Los tópicos de las secciones de comprensión de lectura, escritura y comprensión auditiva son los mismos, pero en el caso de la prueba de

expresión oral el candidato debe desempeñar cinco tareas representativas de cómo un estudiante universitario discutiría con sus colegas y profesores temas referentes a sus estudios.

La evaluación del CAEL es por medio de puntajes llamados Evaluación de Puntajes por Medio de Bandas, los cuales están clasificados desde un puntaje sobre 10 hasta 90, siendo éste último el óptimo y 60 el puntaje mínimo satisfactorio para acceder a una universidad.

Equivalencias de los puntajes obtenidos con CAEL en relación con otras pruebas internacionales de proficiencia. (Véase tabla No. 2).

<b>Tabla No. 2 RELACIÓN DE PUNTAJES ENTRE ALGUNAS PRUEBAS DE PROFICIENCIA INTERNACIONALES</b>		
<b>Examen</b>	<b>Puntaje</b>	<b>Observaciones</b>
<b>FCE</b>	<b>B</b>	<b>La calificación mínima para pasar es C</b>
<b>TOEFL</b>	<b>De 450 a 630</b>	<b>Rangos requeridos por universidades</b>
<b>IELTS</b>	<b>6.5</b>	<b>Prueba con un mínimo de 6 en cada banda de evaluación</b>
<b>CAEL</b>	<b>70</b>	<b>La calificación mínima para pasar es 60</b>
<b>MELAB*</b>	<b>85</b>	<b>Con un puntaje mínimo de 80 en cada sección de la prueba</b>

\* Batería de Evaluación de la Lengua Inglesa de Michigan (MELAB)

El CAEL tiene libre acceso a su política tanto de calificación del examen, como a exámenes ya aplicados y material para preparación de candidatos a fin de que todos los involucrados en la aplicación, calificación y constante revisión y mejoramiento de examen estén enterados de las

políticas vigentes al momento de la aplicación del examen (CAEL Assessment, 2007).

Otras pruebas de proficiencia de la lengua inglesa y de carácter internacional además de las ya mencionadas, son las diseñadas por la Universidad de Michigan a través de su departamento de Servicio de Evaluación de la Lengua Inglesa. Las pruebas son administradas por medio de otro departamento universitario llamado División de Evaluación y Certificación del Instituto de la Lengua Inglesa (ELI).

La Universidad de Michigan tiene una larga historia en el desarrollo de exámenes de inglés como segunda lengua. Dos de los exámenes que más aceptación y uso tienen en el mundo son el Examen para la Certificación en la Competencia del Inglés (ECCE) y el Examen para el Certificado de Proficiencia en Inglés (ECPE). Estos exámenes se aplican en Europa, América del Sur, EE.UU., el Oriente Medio y Asia. En el caso del examen ECCE se utiliza en escuelas para que los alumnos certifiquen sus conocimientos de la lengua inglesa y en el caso de un ambiente laboral se utiliza para evaluar empleados, ofertas de trabajo, etc.

Cada examen ha sido diseñado para candidatos que tengan un nivel alto intermedio de proficiencia en inglés por lo que la prueba tiene cinco secciones. El uso y conocimiento de gramática se prueba por medio de 150 reactivos, la habilidad auditiva con grabaciones seguidas de preguntas relacionadas con lo escuchado, la producción escrita puede

consistir en la redacción de un ensayo o una carta, la lectura está dividida en tres partes, las cuales prueban las diferentes habilidades de lectura que debe conocer el candidato y finalmente en el caso de la expresión oral el candidato tiene una entrevista con el examinador quien utiliza ilustraciones para estimular la interacción con el candidato.

La prueba para el nivel avanzado de proficiencia, ECPE, es la utilizada como requisito para ingresar a una universidad principalmente pero también se usa como certificación de un nivel avanzado de inglés que puede permitir la promoción en un empleo o bien la contratación laboral. Los resultados del ECPE se pueden convertir al puntaje de la Batería de Evaluación de la Lengua Inglesa de Michigan, MELAB, la cual es una prueba muy utilizada por las universidades americanas, canadienses, británicas o bien en países donde la lengua de instrucción no es el inglés. La MELAB es reconocida por sus evaluaciones y certificación de conocimientos ya que les permite a los candidatos obtener un empleo, acceso a una institución educativa o bien obtener beneficios como capacitación para el trabajo, etc. El organismo que administra esta prueba, ELI, ha estado trabajando a fin de incorporar a su sistema de evaluación, los fundamentos de la metodología, de la enseñanza, del aprendizaje y de la evaluación del CEF con el objeto de que sus pruebas tengan los mismos parámetros logrando hasta ahora que se les clasifique con un nivel B2 para la prueba ECCE y la ECPE con un nivel C2,

teniendo de esta manera un reconocimiento en la Comunidad Europea. La comparación entre puntajes del MELAB y su relación con otras pruebas se mostró en la tabla número 2 (The Michigan English Tests, 2007).

Existen otras pruebas de proficiencia que también son utilizadas por universidades, escuelas, departamentos de inmigración o bien para cuestiones laborales, pero se han mencionado las que hasta hoy son más comunes y aceptadas internacionalmente.

#### **4. FCE Y TOEFL – PRUEBAS INTERNACIONALES DE PROFICIENCIA DE LENGUA**

En el capítulo anterior se habló de pruebas de proficiencia de inglés mundialmente aceptadas y reconocidas como instrumentos de medición de proficiencia confiables para fines de admisión en universidades, contratación de personal, inmigración, etc. Ahora se hablará de una de ellas, el First Certificate in English, FCE, de la Universidad de Cambridge y de su organización llamada Cambridge English for Speakers of Other Languages Examinations, Cambridge ESOL Examinations, quien está a cargo del desarrollo, administración y manejo de exámenes de proficiencia de la lengua inglesa que son presentados varias veces al año, por un número aproximado de más de un millón de candidatos, en un número cercano a 1,900 centros para la aplicación de exámenes Cambridge ESOL.

## **4.1 FUNDAMENTOS TEÓRICOS DEL FCE**

La Universidad de Cambridge ha tenido ingerencia en la educación y el diseño de la curricula a nivel del Reino Unido y su interrelación con escuelas, profesores, publicaciones y los departamentos de exámenes, ha permitido que haya un reforzamiento de estándares para la educación de manera constante. Dicha universidad cuenta con una institución llamada Cambridge ESOL Examinations que forma parte de la University of Cambridge Local Examinations Syndicate, UCLES ESOL, la cual ha ofrecido un extenso rango de exámenes, certificados y diplomas a estudiantes del idioma inglés desde 1913. Estos exámenes son precisamente para candidatos que no son nativo hablantes del inglés, son tomados a nivel mundial y por lo tanto tienen un reconocimiento que permite a quien los obtenga tener acceso a instituciones educativas o laborales de diferentes países. En este último aspecto, el FCE es indicativo de una proficiencia suficiente para desempeñarse en áreas administrativas, secretariales, del turismo, etc. Otro aspecto interesante que agregar es que el FCE fue diseñado y desarrollado de acuerdo con la tradición del sistema de exámenes británicos, el cual se fundamenta en el juicio experto y la experiencia institucional para la producción, puntaje y establecimiento de los criterios que se utilizan para interpretar los resultados de la prueba a evaluar, (Bachman, Davidson, Ryan y Choi 1995). Adicionalmente, al ser fundador de la CEF, y socio de la ALTE, Cambridge ESOL Examinations también sigue los parámetros

establecidos por estas autoridades para el diseño, elaboración y evaluación de los exámenes. Cambridge ESOL Examinations tiene un programa permanente de investigación y evaluación así como un continuo monitoreo del sistema de calificación. También se sigue un riguroso sistema de procedimientos para la producción y pre-prueba (pre-test) de las diferentes partes de la prueba con el objetivo de asegurar que las pruebas tengan tanto confiabilidad como validez. El FCE, es un examen muy popular que corresponde a un nivel B2 Vantage de acuerdo con la escala de la ALTE, es decir un nivel intermedio.

El FCE es tomado por alrededor de 270,000 candidatos en más de 100 países cada año. Este examen ha sido diseñado para que sin importar la nacionalidad o género como tampoco la lengua madre del candidato, se obtenga un resultado representativo de su proficiencia en la lengua inglesa. El examen, como todos los que produce la Cambridge ESOL Examinations, está respaldado por un equipo de investigación que se encuentra en el Reino Unido y cuyo programa de pre-prueba permite tener confiabilidad en las tareas y cuestionarios contenidos en los exámenes. Existe un programa permanente de investigación y validación que cubre tres áreas principalmente:

- 1) Análisis operacionales de rutina:** Es decir que se trata de un análisis referente a la administración, producción y conducción de

los exámenes así como a la calificación y la evaluación posterior al examen.

- 2)** Investigación instrumental: Ésta se lleva a cabo por medio de proyectos en pequeña escala que se han diseñado para informar cómo funciona el sistema operativo de los exámenes.
  
- 3)** Proyectos de investigación: Se manejan proyectos de investigación con objetivos a largo plazo, particularmente aquellos que tenga relevancia en la evaluación de la lengua con respecto a los objetivos de Cambridge ESOL.

El programa de investigación también tiene como objetivo asegurar que todos los exámenes cumplen con los estándares relacionados con los siguientes aspectos:

- a)** Validez: hasta qué punto los puntajes obtenidos reflejan la habilidad examinada
- b)** Confiabilidad: hasta que punto los resultados son consistentes, precisos y por tanto confiables.
- c)** Impacto: El efecto que la prueba tiene en los candidatos y otros usuarios.

- d) Factibilidad: hasta qué punto la prueba es práctica en términos de los recursos requeridos.

Los resultados de estas investigaciones son dados a conocer por medio de una publicación mensual que se distribuye en todo el mundo así como por vía electrónica. En dicha información se incluyen estadísticos y los análisis de datos pertinentes que el departamento de investigación ha obtenido por medio de su programa de investigación.

La Universidad de Cambridge ha estado en constante innovación. Así en 1989 se creó un organismo pionero en su tipo en el Reino Unido, llamado Grupo de Investigación y Validación con el fin de establecer un programa de validación con una agenda de investigación enfocada en los exámenes de ESOL. Este grupo se creó como resultado de una investigación realizada entre 1987 y 1989 por Lyle Bachman conocida como The Cambridge – TOEFL Comparability Study y cuyos resultados originaron muchos proyectos. El grupo de investigación mencionado ha crecido para responder a las demandas surgidas de las revisiones que se hacen a las pruebas existentes así como para el desarrollo de nuevas pruebas. Actualmente dicho organismo ofrece un riguroso control de calidad de los exámenes Cambridge ESOL Examinations en cada etapa del proceso de evaluación por medio del establecimiento e implemento de

procedimientos estandarizados, los cuales son evaluados y mejorados desde el punto de vista teórico, técnico y comercial.

## 4.2 Sistema de Evaluación del FCE

El sistema de evaluación de pruebas adoptado por Cambridge ESOL Examinations es el establecido por la CEF y ALTE, es decir que un candidato es evaluado en base a su desempeño general en toda la prueba, y no en base a sólo una de sus partes. Los resultados son expresados en letras. A, B, y C como aprobatorias, El certificado obtenido incluye también las leyendas Excepcional, Good (Bueno) Borderline (En el límite) y Weak (Débil) para cada prueba,. Desafortunadamente, la información disponible no incluye en detalle y de manera precisa, cómo se relacionan los puntajes en todos sus rangos, con las letras asignadas., (véase tabla número 3))

Tabla No. 3 ESCALA APROXIMADA DE PUNTAJE APROBATORIO DE LA PRUEBA FCE	
Letra	Puntajes
A	80% al 100%
B	79% al 61%
C	60%

Además del certificado expedido por Cambridge ESOL Examinations, el Consejo Británico expide un documento en el L1 del candidato en el que se especifica lo que el sujeto puede hacer con el idioma inglés en cada área examinada, (FCE Handbook for teachers. University of Cambridge ESOL Examinations, UCLES 2005. Cambridge.UK).

### **4.3 ELABORACIÓN DE LOS COMPONENTES DEL EXAMEN**

El desarrollo de pruebas que son imparciales y accesibles para todos los candidatos ha sido el común denominador de los exámenes de Cambridge ESOL Examinations a nivel mundial. A continuación se presenta la información relativa a las diferentes partes de la prueba, número de ítems y el tiempo destinado para contestar cada sección., (véase cuadro No. 2)

No todos los ítems de las diferentes secciones del examen tienen un mismo valor, y cada sección del examen vale 40 puntos independientemente del número de ítems que contenga, UCLES ESOL Examinations. El examen consiste en cinco secciones y suman en su conjunto un total de 200 puntos.

<b>Cuadro N o. 2 FORMATO DEL FIRST CERTIFICATE IN ENGLISH</b>		
<b>Sección</b>	<b>Número de ítems</b>	<b>Tiempo</b>
<b>Comprensión de Lectura</b>		
Parte 1: Parear textos con encabezados	6 ó 7	
Parte 2: Opción múltiple		
Objetivo: Encontrar información en detalle	7 ó 8	
Parte 3: Texto con espacios en blanco	6 ó 7	
Parte 4: Pareo múltiple de datos y opción múltiple	13 ó 15	
<b>Total</b>	<b>32 ó 37</b>	<b>1 hora 15 minutos</b>
<b>Escritura</b>		
	<b>Número de Ítems</b>	<b>Tiempo</b>
Parte 1: Escritura de una carta en la que solicita o bien ofrece información (estilo formal / informal).	120 a 180 palabras	
Parte 2: a) Escritura de un documento basado en una situación específica en uno de estos estilos: un reporte, una carta solicitando un empleo, etc., una carta informal, una historia corta, una descripción, una narración, un artículo. b) Un artículo, un reporte, una composición, una historia corta, una narración, una descripción, El tópico para cada estilo a desarrollar es asignado en el examen c) Escritura de un documento en alguno de los estilos mencionados, basado en la lectura opcional de cinco libros asignados por Cambridge ESOL Examinations.	120 a 180 palabras	
<b>Total</b>	<b>Dos documentos de 120 a 180 Palabras en cada documento</b>	<b>1 hora 30 minutos</b>
<b>Gramática y Vocabulario</b>		
	<b>Número de ítems</b>	<b>Tiempo</b>
Parte 1: Ejercicio de opción múltiple basado en vocabulario	15	
Parte 2: Ejercicio con espacios en blanco que deben ser completados con una palabra que complementa un texto.	15	
Parte 3: Ejercicio para evaluar gramática y vocabulario al parafrasear oraciones	10	
Parte 4: Corrección de errores	15	
Parte 5: Formación de palabras	10 a 12	
<b>Total</b>	<b>65 a 67</b>	<b>1 hora 15 minutos</b>
<b>Comprensión Auditiva</b>		
	<b>Número de ítems</b>	<b>Tiempo</b>
Parte 1: Ejercicio de opción múltiple para identificar la información más importante, etc.	8	
Parte 2: Ejercicio para tomar notas	10	
Parte 3: Pareo de información	5	
Parte 4: Ejercicio con dos o 3 posibles respuestas a seleccionar	7	
<b>Total</b>	<b>30</b>	<b>40 minutos aprox.</b>
<b>Expresión Oral</b>		
	<b>Número de ítems</b>	<b>Tiempo</b>
Parte 1: Entrevista a los dos o tres candidatos		3 minutos
Parte 2: Descripción, comparación y contraste de fotografías		4 minutos
Parte 3: Discusión, negociación		3 minutos
Parte 4: Intercambio de información, opinión, desarrollo del tópico discutido previamente por los candidatos.		4 minutos
<b>Total</b>	<b>N/A</b>	<b>14 minutos</b>

El FCE tiene tareas a resolver que son muy similares a situaciones de la vida real, ya que se utiliza material auténtico en su elaboración, razón por la cual toma varios años tener la prueba lista para su aplicación. En el equipo que elabora el examen existen personas llamadas presidentes, que se encargan de vigilar el desarrollo de cada parte del examen. Las personas encargadas de la redacción del examen seleccionan el material auténtico de libros, periódicos, revistas, folletos, anuncios transmitidos por la radio, entre otras fuentes. Las fotos que se seleccionan para la sesión de expresión oral y para el examen de escritura son elaboradas por este grupo de redactores los cuales son maestros practicantes o bien consultores. El trabajo de estos profesionales también incluye la elaboración de guías detalladas que les sirven para la selección de material y la redacción de las preguntas a utilizar en el examen. Dentro de los factores a considerar en esas guías está el tipo de lenguaje, la complejidad lingüística así como las consideraciones de aspecto cultural.

Una vez que se ha seleccionado el material escrito, visual y de audio y las preguntas referentes a dicho material se han elaborado, se turna a un proceso de dos etapas de revisión para su edición. Dicha edición se realiza en equipo y tanto el lenguaje como los temas de la prueba son examinados en detalle para asegurarse de que tienen la relevancia necesaria y que prueba el lenguaje tal y como lo marcan las

especificaciones para la elaboración de pruebas de Cambridge ESOL Examinations. Una vez que los cuestionarios del examen han sido corregidos y/o aceptados se incorporan a un sofisticado sistema de software que ha sido diseñado específicamente para Cambridge ESOL Examinations y que permite la rápida elaboración de exámenes, pero que al mismo tiempo respalda todas las actividades asociadas al proceso de producción de pruebas. Cuando el encargado de los exámenes en proceso, y el presidente están de acuerdo en que el material editado es correcto en todas las tareas incluidas en el examen, se hace un pre-test el cual implica una muestra de más de 100 candidatos en centros donde se aplican los exámenes Cambridge o bien en escuelas, donde se aplican las pruebas de lectura, gramática y de comprensión auditiva ya que estas pueden ser analizadas desde el punto de vista estadístico.

Todas las pruebas para pre-test son aplicadas en todo el mundo ya que el objetivo de este proceso es establecer si el material utilizado tiene una calidad adecuada para ser utilizado en un examen oficial. También se busca encontrar características que puedan ser medidas y que permitan la elaboración de pruebas que tengan una dificultad comparativamente igual a las utilizadas en el pre-test.

En el caso de las pruebas de escritura y de expresión oral, el pre-test se lleva a cabo igualmente en centros de aplicación de exámenes Cambridge o escuelas, donde participan pequeños grupos de candidatos. El objetivo es que las tareas a realizar así como las rúbricas y el material visual

utilizado en la entrevista que se conduce en el examen de habilidad oral, tengan un nivel apropiado de dificultad, sean claros y objetivos. Con respecto al caso de la escritura el número de candidatos que participan en el pre-test es de 15 a 20 en igualdad de circunstancias en cuanto al tipo de lugar donde se aplica la prueba.

El conjunto de encargados de elaborar las pruebas FCE deben dar su aprobación a la prueba. Posteriormente, ésta es analizada por una persona llamada calificador, quien no ha tenido contacto previo con la prueba, y la analiza principalmente en cuanto a su contenido. Después de esta etapa se analiza la tipografía del examen y se llevan a cabo posteriores análisis de redacción, pero en cada etapa el encargado de los exámenes revisa nuevamente cada examen cuidadosamente y por completo y en detalle. En la etapa final el examen es leído y analizado por dos personas externas que además contestan el examen para asegurarse de que las respuestas son las esperadas en las pruebas de comprensión de lectura, gramática y de comprensión auditiva. En todos los procesos de revisión el personal encargado recibe una lista de aspectos a verificar e instrucciones precisas de lo que se espera que hagan en su análisis. Una última revisión tanto de redacción, tipografía, contenido como de la respuestas esperadas en el examen se realiza antes de que el examen se envíe a impresión y una vez impreso se vuelve a revisar para asegurarse de que tiene la calidad y cumple con las especificaciones de Cambridge

ESOL Examinations para poder utilizarlo en un examen oficial en todo el mundo (Cambridge ESOL – First Certificate Exam specifications, 2007).

### **4.3.1 Características de la Prueba de Comprensión de Lectura**

El examen de comprensión de lectura toma una hora con 15 minutos y el objetivo es evaluar la habilidad del candidato para utilizar diferentes estrategias de lectura a través de las cuatro partes en que consiste el examen. En la primera parte se tiene como objetivo poner a prueba la habilidad del candidato para identificar los puntos más importantes en un texto a nivel de párrafos, debiendo realizar tareas que incluyen la identificación de encabezados o bien oraciones que resumen ideas que deben ser pareadas con un texto. La segunda parte del examen consiste en un ejercicio de opción múltiple donde el candidato debe identificar oraciones que resumen ideas y la comprensión detallada del texto. La tercera parte tiene como objetivo evaluar si el candidato puede identificar como se estructura un texto, para ello se presenta un ejercicio con espacios en blanco y la información que complementa el texto en esos espacios es colocada en párrafos de manera aleatoria, a fin de que el candidato identifique la información que debe completar el texto correctamente. Por último se presenta un ejercicio con características muy similares a las de la primera parte. El valor de los ítems en cada una

de las tareas es muy similar entre sí teniendo un puntaje global de 40 puntos aunque la prueba puede contener de 35 a 37 preguntas, regularmente de la pregunta 1 a la 20 se le asigna un valor de dos puntos por pregunta y al resto un punto por pregunta. Se llega al puntaje de 40 por medio de un ajuste.

### **4.3.2 Características de la Prueba de Escritura**

Un candidato puede desempeñarse mejor cuando las tareas a realizar están relacionadas con sus intereses y su experiencia personal con alguno de los temas contenidos en la prueba. Esta parte del examen toma una hora con 30 minutos, los candidatos deben redactar un documento utilizando un mínimo de 120 y hasta un máximo de 180 palabras. Hay dos tareas a realizar, la primera de ellas es obligatoria la cual regularmente es una carta en la que el candidato deberá solicitar o proporcionar información específica contenida en el examen, o bien redactar una queja dirigida a alguna institución o negocio, etc. En las instrucciones para este primer documento, usualmente se incluye un texto con un dibujo, ilustración o diagrama para facilitar la comprensión de la tarea a realizar. El segundo documento a redactar tiene varios tópicos y estilos a elegir y se puede seleccionar entre elaborar una composición casi siempre dirigida a un profesor como una continuación de una actividad en clase. Se puede redactar un artículo para una revista o un boletín, un

reporte dirigido a un supervisor, jefe, etc., también puede ser una carta solicitando una beca, un empleo, etc. Otra tarea es la redacción de una carta informal dirigida a un amigo por correspondencia o bien la narración de una pequeña historia para una revista.

Por medio de estas dos tareas se espera que el candidato demuestre un buen desarrollo y organización de ideas y el uso efectivo de la lengua escrita utilizando los elementos del discurso adecuados al contexto que el examen especifica. Existe una tercera tarea opcional para redactar un documento: el Cambridge ESOL Examinations asigna anualmente la lectura de cinco libros de la literatura clásica de la lengua inglesa, para aquellos candidatos que durante su preparación para el examen desean leer, discutir y escribir acerca de los temas e historias de los libros leídos. Este documento sería en lugar de elegir la segunda tarea descrita anteriormente, teniendo así que producir únicamente dos documentos de 120 a 180 palabras cada uno.

La evaluación de esta sección se hace con el sistema interjuez: la composición es calificada por un profesor capacitado para este tipo de examen y su evaluación es enviada a la base de datos sin que el segundo profesor la conozca antes o después de que sea calificado por segunda vez. El examen se califica utilizando una escala específica dividida en cinco bandas que van del cero al cinco y que se ajustan a un puntaje global de 40 puntos. Dichas bandas contienen especificaciones que les

permiten a los examinadores calificar en detalle el documento que se les asigna ya que se toman en cuenta estos aspectos:

- a) Contenido del documento relacionado con la tarea especificada
- b) Organización y Cohesión
- c) Registro y formato del documento
- d) Precisión en el uso del lenguaje en cuanto a ortografía, vocabulario, gramática.
- e) El lector receptor del documento recibiría un documento con información clara, concisa, etc.

Finalmente cada examinador tiene un formato donde especifica su evaluación punto por punto basado en las especificaciones y asigna una banda a su evaluación.

### **4.3.3 Características de la Prueba de Gramática y Vocabulario**

La prueba de gramática y vocabulario se realiza en una hora con 15 minutos. Esta sección tiene cinco partes. La primera es un ejercicio de opción múltiple basado en vocabulario donde se hace énfasis en el uso correcto del vocabulario dentro de un texto que tiene 15 espacios en blanco y uno de ejemplo. La segunda parte es un ejercicio con espacios

en blanco que deben ser completados con una palabra que complementa un texto, el objetivo es evaluar los conocimientos de gramática y vocabulario. En la tercera parte de la prueba se presenta un ejercicio para evaluar el uso correcto de la gramática y vocabulario al tener que re-escribir enunciados que deben ser parafraseados. En la cuarta parte el candidato tiene que identificar errores gramaticales en un texto y finalmente en la quinta parte se pide a los candidatos derivar palabras directamente relacionadas con la gramática y vocabulario utilizado en un texto. El total de preguntas es 65, el valor de las preguntas que van del 1 al 30 y de la 41 a la 65 reciben un punto, de las preguntas 31 a la 40 se les asigna dos puntos, el puntaje global se ajusta según informa Cambridge ESOL por medio de un sofisticado sistema para obtener un puntaje total de 40 (Cambridge ESOL First Certificate Exam Specifications,2007).

#### **4.3.4 Características de la Prueba de Comprensión Auditiva**

Esta sección del examen está dividida en cuatro partes y su duración es de aproximadamente 40 minutos. El objetivo es medir la habilidad del candidato para escuchar información detallada en textos cortos y con una variedad de enfoques. En la primera parte del examen se utiliza el formato de opción múltiple, en la segunda el candidato debe

tomar notas de la información escuchada, ya que se mide su habilidad para identificar vocabulario y frases que le permitan resumir información, así como también producir respuestas que complementen la información contenida en esta parte del examen. La tercera parte no tiene el mismo formato en todos los exámenes, por ejemplo uno de estos formatos consiste en la identificación de información en relación con el contenido de un diálogo entre diferentes participantes por medio del sistema de opción múltiple, otro de ellos consiste en escuchar un texto largo y resumir información. La cuarta y última tarea en esta sección generalmente consiste en un ejercicio de verdadero o falso o bien uno en el que haya tres opciones a seleccionar. El objetivo es evaluar si el candidato puede discriminar la información más importante en una conversación, las funciones utilizadas, etc. Los textos escuchados en las cuatro secciones contienen diferentes tipos de acentos y pronunciación, los temas incluidos en los exámenes son variados y van desde cuestiones culturales, viajes, eventos históricos, costumbres, etc.

El total de puntos en el examen es de 30 puntos y se ajusta para obtener un puntaje de 40.

#### **4.3.5 Características de la Prueba de Expresión Oral**

Para esta sección se requiere de dos examinadores, los cuales son seleccionados y capacitados constantemente, uno que administra

(interlocutor) el examen y otro más que califica (asesor) basándose en una escala específica de bandas del cero al cinco los siguientes aspectos:

- a)** Gramática y vocabulario
- b)** Manejo del discurso
- c)** Pronunciación
- d)** Comunicación interactiva

El interlocutor da una calificación de desempeño global basándose en una escala simplificada que especifica la banda del cero al cinco. El 5 corresponde a un candidato que demuestra tener habilidades interactivas, dar un mensaje claro, resolver la situación que se le pide en una tarea específica, entre otros aspectos. Dichas calificaciones las registra el asesor en un formato y no se discuten ya que se trata de una evaluación interjuez.

La duración del examen es de 14 minutos aproximadamente para dos alumnos, 20 en el caso de un trío, ya que el examen se toma en pareja, pero en caso necesario se aceptan tres candidatos ya que se evita tener una prueba en la que un candidato interactúe directamente con un examinador. En un caso extremo se le pide a uno de los candidatos que ya tomaron el examen que participe en el examen para que el candidato examinado pueda interactuar con alguien más y no con el examinador, debido a que se busca tener un contexto de la vida real en un ambiente relajado, para que el candidato produzca un lenguaje lo más cercano posible al que se utilizaría en una situación cotidiana. Las tareas a

desempeñar son medidas cuidadosamente por el interlocutor y se sigue un riguroso guión de instrucciones que contiene tanto preguntas, como rúbricas y material visual para el examen:

- Primera parte:** Conversación entre el interlocutor y el candidato basada en preguntas específicas en el guión
- Segunda parte:** Los candidatos deben describir y comparar un par de fotografías y expresar su opinión respecto a ellas.
- Tercera parte:** Los candidatos deben conversar para decidir sobre una situación utilizando material visual como guía.
- Cuarta parte:** Discusión basada en la tarea anterior, el candidato expresa su opinión, desacuerdo, etc.

El objetivo de esta prueba es que los candidatos puedan demostrar su habilidad en el uso del lenguaje hablado en un ambiente amigable aunque controlado mientras que los examinadores deben mantener las condiciones favorables para que el arreglo de sillas y mesa de trabajo no interfiera con la comunicación cara a cara entre los candidatos y el interlocutor. En el formato de examen donde se registran las calificaciones se indica con quien interactuó cada candidato, el número de prueba y materiales utilizados además se graban algunas de las entrevistas a fin de

que un examinador externo asignado por Cambridge ESOL verifique la calidad de la entrevista, las calificaciones reportadas y si el examen se condujo de acuerdo a las especificaciones estipuladas.

#### **4. 4 Validez y Confiabilidad en el FCE**

El Cambridge ESOL Examinations tiene un departamento de investigación dedicado a llevar a cabo todo tipo de análisis, comparaciones, investigación, etc., para conocer tanto la validez como la confiabilidad, entre otros aspectos, de las diferentes pruebas que ellos producen y que se aplican en el mundo. Por lo tanto esas investigaciones no se realizan solo en el Reino Unido, sino también en cualquier país donde se apliquen los exámenes de Cambridge ESOL Examinations. Este tipo de investigación es permanente, aún después de haber elaborado sus exámenes siguiendo los estrictos lineamientos de la ALTE y CEF. Adicionalmente, este departamento organiza conferencias sobre evaluación, y todo lo relacionado con la elaboración de pruebas de proficiencia de lengua inglesa principalmente. Con la evidencia obtenida a través de la investigación proponen formas de mejorar las pruebas ya existentes así como la elaboración de nuevos exámenes que respondan a las necesidades académicas de instituciones desde el nivel universitario hasta de educación básica, es decir para niños de primaria. Además de

todo esto, el departamento de investigación promueve la publicación de los avances y estudios sobre evaluación a nivel mundial.

En una de las publicaciones del departamento de investigación, (Research Notes 4 February, 2001) el coordinador del departamento, Neil Jones, afirma que un aspecto que contribuye a la confiabilidad de los exámenes de Cambridge ESOL Examinations, es la metodología tanto para la aplicación de los exámenes como para su calificación. Agregan que otro factor que forma parte de la confiabilidad dentro de todo el sistema de sus exámenes, es la preparación que debe tomar el candidato que desee hacer una de sus pruebas.

Las publicaciones de textos como material de estudio para los exámenes, por parte de editoriales en todo el mundo, tienen el mismo formato según el examen de que se trate y se observan los mismos parámetros de evaluación. El objetivo es que quien se prepare por medio de este tipo de cursos, estudie el inglés para adquirirlo, no para pasar un examen. Con respecto a la validez se cita a Bachman (1990) en esta misma publicación y dicen estar de acuerdo con este autor cuando propone que la confiabilidad sea vista como la acción de minimizar los efectos de los errores en las mediciones, y a la validez como la acción de maximizar los efectos de las habilidades lingüísticas que se desean medir, concluyendo que por lo tanto confiabilidad y validez son complementarios. El departamento de investigación dice estar conforme con estos conceptos, porque son parte del objetivo de Cambridge ESOL

Examinations, el cual es: medir, identificar, estimar y controlar los efectos de los factores que afectan los puntajes de las pruebas, así como producir pruebas que sean confiables y justas.

Aunque no se tiene información respecto a un estudio de la validez concurrente y/o la consistencia interna por medio de una correlación entre TOEFL y FCE, existe una investigación cuyo objetivo principal es la comparabilidad en la medición de las habilidades entre estas dos pruebas, realizado por Bachman, Davidson, Ryan y Choi, (1995). Para tal efecto se estudio la validez de constructo, lo que llevó al análisis de la calidad del contenido de las dos pruebas. Asimismo se investigó la confiabilidad entre las dos pruebas y su consistencia interna, pero este último aspecto no como objetivo principal del estudio sino como un complemento de la investigación misma. Esta investigación se diseñó considerando los países donde cada una de las pruebas es aplicada con mayor frecuencia. Para el TOEFL los países seleccionados fueron Tailandia, Egipto, Japón y Hong Kong. Para el FCE se seleccionaron España, Brasil, Francia y Suiza. El total de sujetos participantes fue de 1,450, el personal que regularmente es responsable de la aplicación de estos exámenes en esos sitios fue designado para aplicar los exámenes que se utilizarían en la investigación. Los sujetos se seleccionaron tomando en cuenta un promedio de edad, preparación para el examen cuestión de estudio, y conocimientos previos de inglés, entre otros aspectos. Se usaron dos versiones diferentes de las

pruebas, a excepción de Expresión Oral del FCE, una por cada dos países. Las pruebas, TOEFL y FCE, se administraron en las mismas fechas en todos los países.

Los investigadores dicen que mucha de la información recopilada fue en referencia a la medición e interpretación de la validez de constructo de las dos pruebas, la cual les fue difícil analizar, pero más y en particular en el FCE, debido que existe una amplia variedad de habilidades implicadas en cada una de las pruebas lo que confirma, según afirman, que la proficiencia en una lengua es un constructo multicomponente, que cuando se mide, se ve afectado en mucho por la metodología de aplicación del instrumento de medición. Los datos obtenidos revelaron que la validez de constructo, en las diferentes secciones del examen del FCE, tiene algunas deficiencias que deben analizarse de manera empírica y estudiar sus efectos en los puntajes. Existen específicamente tres áreas que Bachman et al recomiendan estudiar debido a su complejidad:

- a) El diseño de la prueba
- b) La administración de la prueba
- c) El método de calificación origina variaciones en los puntajes

Con estas recomendaciones los investigadores argumentan que se podrían reducir o bien controlar las fuentes de falta de confiabilidad en las

diferentes áreas del examen ya que no requieren de mayores modificaciones. En cuanto a la confiabilidad del FCE y en relación a la consistencia interna, el estudio indica que estudiando por separado cada prueba, se encontraron los siguientes datos (véase tabla no.4).

<b>Tabla No. 4 REPORTE SOBRE CONFIABILIDAD CONSISTENCIA INTERNA DEL FCE</b>	
<b>Sección de la Prueba</b>	<b>Consistencia Interna</b>
Comprensión de Lectura	Fuerte
Escritura	Baja
Gramática y Vocabulario	Fuerte
Comprensión Auditiva	Baja
Expresión Oral	Baja

Las inconsistencias en cuanto al sistema de calificación principalmente, dicen los investigadores, no les permitió interpretar los puntajes de manera confiable y poder determinar estadísticamente la consistencia interna en todas las secciones del FCE.

## 4. 5 FUNDAMENTOS TEÓRICOS DE TOEFL

El examen TOEFL, The Test of English as a Foreign Language, es un examen utilizado para evaluar la proficiencia de la lengua inglesa de personas no nativo hablantes del inglés. En sus orígenes, este examen fue desarrollado para medir la habilidad de estudiantes universitarios no nativo hablantes del inglés, para usar y comprender la lengua inglesa tanto en la expresión oral, escrita y comprensión auditiva en contextos de enseñanza a nivel universitario. Esta prueba es considerada como el prototipo de las pruebas de lengua clasificadas como “sicométricas/estructuralistas” tradición que se ha respetado en la estructuración de los componentes de prueba escrita y de expresión oral (Bachman Davidson, Ryan, Choi. 1995). Actualmente el examen se ha diversificado y se utiliza para obtener becas, obtener un empleo o bien una promoción en el mundo. El Educational Testing Service, ETS, en Princeton, Nueva Jersey, EUA., es la institución que ha desarrollado, maneja y administra varios exámenes de proficiencia del inglés a nivel internacional, siendo el TOEFL uno de ellos. Esta organización fue fundada en 1947 y en la actualidad tiene presencia en 180 países, en más de 9,000 locaciones y con una aplicación anual de más de 24 millones de pruebas en el mundo. El ETS se dice ser una organización no lucrativa cuya misión es lograr el desarrollo y avance en la calidad e igualdad en la educación, por medio de exámenes que aporten una evaluación precisa,

justa y válida. Esta postura se basa en los trabajos de investigación en el campo de la evaluación que el ETS lleva a cabo a nivel internacional.

El objetivo del examen TOEFL es ofrecer a las instituciones educativas, empresas y gobiernos en todo el mundo, un producto confiable y válido basado en las necesidades tanto de los usuarios como de los candidatos. Dichas necesidades son investigadas y analizadas y conjuntamente se desarrollan principios para comprender los modelos psicométricos relacionados con la prueba, para que de acuerdo a la complejidad requerida se elabore el examen. Otro objetivo que el ETS tiene con respecto a sus pruebas, y en este caso TOEFL, es ofrecer un producto que sea resultado del estudio cuidadoso de los constructos con los que fue elaborado, teniendo una clara finalidad cada uno de ellos dentro del examen. Es también importante la forma de medición y la definición de lo que miden los constructos en la prueba, por lo que son estudiados, puestos en práctica y aprobados contra los estándares de calidad establecidos por el ETS. Existen dos tipos de constructos que se toman en cuenta para elaborar la prueba TOEFL, los de tipo cognitivo, que son: el pensamiento crítico, las habilidades comunicativas y la inteligencia emocional. El otro tipo de constructo es el no cognitivo es decir, la persistencia, formalidad y/o seguridad y el trabajo en equipo. El problema desde el punto de vista científico lo constituye el encontrar constructos que representen éxito en la educación, el desarrollo de medios para medir

los nuevos constructos y la exploración de diversos usos para esos nuevos constructos.

La elaboración de pruebas TOEFL es constante y para ello tanto los investigadores encargados de estudiar el comportamiento de la prueba para conocer cuestiones como validez o confiabilidad, por ejemplo, como el departamento de Desarrollo de la Evaluación, son las instancias que consulta la ETS cuando se planea y se diseña una prueba. Los clientes del ETS tienen una relación cercana con el departamento de Desarrollo de la Evaluación, la finalidad es que la información recabada por este departamento con respecto a los requerimientos, problemática o necesidades de los clientes, sea considerada para desarrollar y finalizar las especificaciones que requiere una prueba TOEFL. Dentro de los objetivos del ETS en el proceso de elaboración de la prueba, es el de asegurarse de que el material que se utiliza en este examen sea similar al que un estudiante encontraría en un contexto universitario. Un objetivo más, es la selección de contextos donde por medio del sistema de pre-prueba se obtengan resultados que les permitan definir los puntajes que se utilizarán en la evaluación de las diferentes secciones de la prueba. Este proceso consta de cinco etapas las cuales son:

- 1) Planeación y diseño de la evaluación
- 2) Escritura y revisión de los cuestionarios del examen  
(Esta etapa incluye preguntas en el campo y de pre-prueba)

- 3) La recolección de cuestionarios para la elaboración de pruebas o bien para los centros de concentración de pruebas
- 4) Publicación de las pruebas
- 5) Administración y calificación de las pruebas

#### 4.5.1 Sistema de Evaluación del TOEFL

El puntaje obtenido por un candidato no corresponde al porcentaje de las respuestas correctas del examen. El resultado se obtiene por medio de sofisticados procedimientos estadísticos que se usan para convertir esas respuestas correctas en un puntaje que se iguala a los resultados obtenidos por otros candidatos con habilidades similares. Es decir, que si una prueba en particular resultara tener un mayor rango de dificultad, el puntaje convertido del candidato reflejará un resultado similar al de otro candidato que tomó una prueba con menor rango de dificultad, (Ver tabla No. 5).

<b>Tabla No. 5 ESCALA DE PUNTAJE DE LA PRUEBA TOEFL</b>	
<b>Sección</b>	<b>Puntajes</b>
Comprensión Auditiva	De 31 a 68
Expresión Escrita y Gramática	De 31 a 68
Vocabulario y Comprensión de Lectura	De 31 a 67
Prueba de Expresión Escrita	De 1 a 6
Puntaje total	De 310 a 677

Cada institución define el puntaje requerido para que los alumnos puedan acceder a su inscripción, actualmente en promedio para acceder a una universidad en Canadá o Estados Unidos, se requieren 600 puntos. Sin

embargo la vigencia del puntaje obtenido por un candidato es de hasta dos años, ya que el ETS considera que la proficiencia de un candidato puede variar considerablemente en un período bastante corto posterior al reporte de los resultados. Por lo tanto si un candidato desea que su puntaje sea renovado debe tomar el examen nuevamente si ya han transcurrido dos años después de que tomó la prueba.

Para conocer la equivalencia de los rangos de los puntajes de sus pruebas con los niveles de proficiencia del CEF, el ETS llevó a cabo una investigación. Los resultados para el TOEFL de acuerdo a los niveles de proficiencia definidos por el CEF, van del Usuario Independiente al Usuario Proficiente, (véase tabla No. 6)

Tabla No. 6 EQUIVALENCIA DE NIVELES DE PROFICIENCIA DEL TOEFL Y EL CEF			
Sección	Rango del Puntaje	Puntaje B1 Mínimo	C1 Puntaje Mínimo
TOEFL Total	310 - 677	457	560
Comp. Auditiva	31 - 68	46	56
Comp. De Lectura	31 - 67	46	56
Gramática	31 - 68	45	56

(Equivalencias del ETS: Mapping Test Scores onto the Common European Framework)

## 4.5.2 ELABORACIÓN DE LOS COMPONENTES DEL EXAMEN

Para la elaboración de pruebas el ETS tiene un Departamento de Desarrollo de la Evaluación. La producción de exámenes es de más de 180 pruebas de diferentes tipos. El TOEFL es uno de esos exámenes el cual es ofrecido en dos modalidades: la tradicional prueba en papel para ser respondida con lápiz y la prueba por medio de Internet. Esta última

prueba es la que actualmente se está aplicando con más frecuencia prácticamente en todo el mundo. A continuación se hablará de la prueba en papel ya que fue la que se utilizó en este trabajo.

Este examen consta de tres secciones: la prueba de comprensión auditiva, la prueba de estructura y expresión escrita y la de vocabulario y comprensión de lectura. También existe la prueba de escritura en inglés, pero ésta solo se aplica cinco veces al año ya que pocas instituciones la requieren, por lo que es una sección que no se incluye de manera regular en el examen.

A continuación se presentan los componentes del TOEFL, (véase tabla No. 7)).

<b>Tabla No. 7</b>		
<b>FORMATO DE TOEFL</b>		
<b>Sección</b>	<b>Número de ítems</b>	<b>Tiempo</b>
<b>Comprensión Auditiva</b>		
Parte A: identificación de enunciados similares.	20	
Parte B: preguntas basadas en conversaciones cortas.	15	
Parte C: preguntas basadas en mini exposiciones y conversaciones cortas.	15	
<b>Total</b>	<b>50</b>	<b>30 – 40 minutos</b>
<b>Expresión Escrita y Gramática</b>		
Identificación de complementos correctos.	15	
Identificación de frases y palabras incorrectas.	25	
<b>Total</b>	<b>40</b>	<b>25 minutos</b>
<b>Vocabulario y Comprensión de Lectura</b>		
<b>Número de ítems</b>		
Identificación de sinónimos, frases o palabras con significado similar.	30	
Lectura de pasajes	30	
<b>Total</b>	<b>60</b>	<b>45 minutos</b>
<b>Prueba de Expresión Escrita en Inglés*</b>	<b>250 a 300 palabras</b>	<b>30 minutos</b>

\* se ofrece cinco veces al año

### **4.5.3 Características de la Prueba de Comprensión**

#### **Auditiva**

Esta sección consta de 50 ítems y tiene una duración de 30 a 40 minutos. El objetivo es que el candidato demuestre su habilidad para discriminar información concisa en diversos contextos, el sistema de respuesta es de opción múltiple. Toda la información se presenta una sola vez, en las tres secciones, no hay repetición. En la primera parte, el candidato debe discriminar información en un ejercicio escrito que consiste en la identificación de enunciados, los cuales tienen similitudes entre sí basados en la información escuchada. La segunda parte consiste en preguntas de comprensión basadas en conversaciones de corta duración y donde una tercera persona hace una pregunta acerca de lo expresado en dicha conversación. En la tercera parte las conversaciones son más largas y además incluye pequeñas exposiciones, Inmediatamente después de escuchar la conversación o exposición se escuchan preguntas las cuales tienen como finalidad evaluar si el candidato puede no solo comprender sino también inferir información hablada.

#### **4.5.4 Características de la Prueba de Expresión Escrita y Gramática**

Esta sección tiene una duración de 25 minutos y tiene 40 ítems. El objetivo en esta sección es que el candidato demuestre su capacidad para identificar los elementos del discurso que complementan adecuadamente un texto así como una serie de frases y vocabulario incorrecto que debe ser identificado en un contexto. En la primera parte el candidato debe completar enunciados que tienen espacios en blanco con el fin de evaluar su conocimiento de un amplio rango de estructuras gramaticales. En la segunda parte el candidato debe identificar las partes de los enunciados que contengan errores.

#### **4.5.5 Características de la Prueba de Vocabulario y Comprensión de Lectura**

El objetivo de esta sección es medir la comprensión de inglés escrito y el formato de respuesta es de opción múltiple también. La prueba consiste en 60 ítems que deben ser resueltos en 45 minutos. En la primera parte el candidato debe identificar sinónimos o palabras muy parecidas en su significado. Después de leer los enunciados del ejercicio los cuales tienen una palabra o una frase subrayada, se presentan cuatro palabras a elegir y que tienen el mismo significado o al menos muy similar a la palabra o frase subrayada en el enunciado. En la segunda parte el

candidato debe leer cuatro pasajes y contestar preguntas de comprensión de cada uno de ellos.

#### **4.5.6 Características de la Prueba Escrita (ensayo)**

Este examen se ofrece cinco veces al año y solamente en caso de que la institución requiera que el candidato lo presente. El candidato resuelve esta sección desarrollando un documento de 250 a 300 palabras en forma de ensayo en un período de 30 minutos. En esta sección el candidato deberá utilizar cualquiera de estos aspectos, de acuerdo al tema especificado en la prueba, y que puede ser: expresar opiniones con argumentos que las respalden, comparación y contraste de información o puntos de vista, comparar y contrastar un tópico, presentar un argumento y persuadir al lector del documento acerca de un tema específico. Para calificar esta prueba se tiene un sistema diferente de evaluación ya que se trata de un ensayo, razón por la que se utiliza el sistema interjuez, otorgando cada evaluador una calificación de manera independiente y secreta con una escala del uno al seis. El ensayo se califica de manera holística, como la llama el ETS, (véase apéndice 1, Scoring Guide for TOEFL) siendo seis el puntaje más alto. Se toma en cuenta la organización y presentación de las ideas, lo bien que se expresen los argumentos y se responda a la pregunta del examen. Los errores ortográficos, gramaticales o de puntuación no son tomados en cuenta en tanto no oscurezcan la claridad del ensayo. Cuando un candidato obtiene

un puntaje de seis, significa que se trata de alguien con muy buena habilidad para expresarse de manera escrita, mientras que cinco representa una habilidad promedio y un cuatro una habilidad muy baja en esta prueba. Por supuesto que un puntaje inferior no es satisfactorio pero no afecta el puntaje general de TOEFL, ya que esta sección se reporta por separado, ETS (2006 Research Agenda).

#### **4.6 Validez y Confiabilidad en el TOEFL**

Para asegurar la validez y confiabilidad de las pruebas, el ETS constantemente investiga, estudia y compara sus pruebas con otras de proficiencia, en el caso de TOEFL se lleva a cabo la investigación de las diferentes secciones de la prueba TOEFL para conocer su validez y confiabilidad. El departamento encargado de esta tarea es el Centro de Investigación y Política de Evaluación, todo inicia desde la elaboración de los ítems de las secciones de la prueba TOEFL, los cuales son revisados y evaluados, en más de 25 ocasiones, hasta lograr el cumplimiento de las especificaciones y estándares del ETS. Posteriormente, los puntajes obtenidos por los candidatos son analizados estadísticamente para conocer su validez y confiabilidad ítem por ítem, determinando así la precisión también, de las respuestas de los ítems en los modelos utilizados en las pruebas.

Existe un comité de examinadores que está integrado por diez especialistas en lingüística, profesores de inglés como segunda lengua y

especialistas en evaluación. Este personal se rota constantemente para asegurar la introducción de nuevas ideas y filosofías relacionadas con la evaluación y enseñanza del inglés como segunda lengua, tomando en cuenta siempre las actualizaciones en las tendencias y metodología en estos campos. La principal responsabilidad de este comité es establecer y revisar todas las políticas, lineamientos y especificaciones que conllevan la elaboración de la prueba, las habilidades a medir, el tipo de preguntas a incluir y el aspecto cultural implícito en el TOEFL. En suma, el objetivo es producir un instrumento que sea válido en la medición de la proficiencia de la lengua inglesa, (Des Brisay, 1989).

En el ETS se cuenta con un Departamento de Investigación y Desarrollo en donde personal altamente calificado académicamente, se encarga de la evaluación y la investigación de las pruebas que administra la institución. Para ello cuentan con centros de investigación que estudian los puntajes contra las pruebas por medio de análisis estadísticos, psicometría, la investigación del proceso de la enseñanza-aprendizaje, el desarrollo de nuevos constructos, la transferencia de la tecnología y la innovación del sistema de evaluación así como investigación de la validez de las pruebas. En su página electrónica el ETS publica sus investigaciones y hace del conocimiento público la información referente al comportamiento de sus diversas pruebas según lo reporta su Departamento de Investigación y Desarrollo.

Dentro de los resultados reportados en su investigación, Bachman et al informan que el TOEFL tiene un diseño, una construcción, una administración, un sistema operacional y de calificación, menos complejo que el FCE. Todo esto les permitió concluir a los investigadores de referencia que el TOEFL es un instrumento de medición con muy pocas fuentes potenciales de error. Estas características también les facilitaron estimar la consistencia interna de la prueba y reportan que encontraron una consistencia uniformemente alta en todas sus partes.

## **5. METODOLOGÍA**

Las pruebas utilizadas para este trabajo son las que publican las editoriales para preparar candidatos que desean tomar estos exámenes. Por lo tanto para este ejercicio se utilizó una prueba FCE que fue aplicada con anterioridad a nivel mundial.

El Cambridge ESOL Examinations autoriza para su publicación muchos de las pruebas aplicadas ya que posteriormente se consideran pruebas de práctica, debido a que no habrá replica de los ejercicios en dichas pruebas (véase apéndice 3). En el caso de TOEFL la editorial que publica las pruebas no hace mención con respecto a su uso previo como pruebas oficiales, pero sigue todos los lineamientos establecidos por el ETS en cuanto a su presentación, instrucciones para los candidatos, tiempo de aplicación, etc., (véase apéndice 4).

## **5. 1 Características de los Participantes**

Para llevar a cabo este trabajo se buscó una preparatoria que tuviera un programa de inglés enfocado en la preparación de alumnos para los exámenes de la Universidad de Cambridge. El programa de inglés de la institución elegida consiste en clases diarias de 90 minutos impartidas por niveles de conocimientos del idioma: básico, intermedio y avanzado. El grupo seleccionado corresponde al nivel avanzado, el cual en esta escuela regularmente recibe preparación para tomar el examen FCE. Los sujetos de dicho grupo estudiaron inglés durante toda la preparatoria con ese programa y los alumnos manifestaron su interés por participar en esta investigación para conocer su desempeño en las dos pruebas, el TOEFL y el FCE, ya que nunca habían recibido ni información ni capacitación para tomar el examen TOEFL y les pareció por lo tanto, que el ejercicio les daría una panorámica diferente en cuanto a puntajes con respecto a su preparación para tomar un examen de proficiencia de

inglés. Dichos sujetos estudiaron inglés como materia obligatoria durante su educación primaria, secundaria y preparatoria y en particular tomaron preparación intensiva para presentar los exámenes de la Universidad de Cambridge, durante la secundaria el examen PET y en la preparatoria el FCE.

El grupo de sujetos participantes estaba conformado por alumnos de clase media y media alta que cursaban el quinto semestre de preparatoria de la escuela particular Liceo Consuelo Rubio de Ruiz ubicada en la ciudad de Querétaro. Dicha institución no es considerada bilingüe pero la enseñanza del inglés es curricular y se imparte desde pre-escolar hasta estudios superiores, razón por la que los alumnos toman clases de inglés con la finalidad de obtener la certificación de sus conocimientos, por medio de los exámenes Cambridge, desde el sexto de primaria. Los participantes en esta investigación eran 12 mujeres y 11 hombres con una edad promedio de 17.5 para el sexo femenino y 17.8 para el sexo masculino. Todos ellos estudiaron inglés desde la escuela primaria, 8 de los sujetos asistieron a escuelas bilingües y habían cursado la preparatoria en la misma escuela con el mismo sistema de enseñanza de inglés por cinco semestres consecutivos. Dos de los participantes estudiaron en Estados Unidos, uno de ellos del sexo femenino quien cursó cinco años de la instrucción primaria y dos años de secundaria y el otro sujeto, del sexo masculino, cursó un año de secundaria en ese país.

Los otros 4 sujetos viajaron a Canadá y Estados Unidos para tomar cursos de verano donde interactuaron con nativo hablantes del inglés dentro de las instituciones donde estudiaron y al hospedarse con familias de nativo hablantes del inglés (véase apéndice 5).

Los profesores de inglés de la preparatoria, no son nativo hablantes de inglés y los textos utilizados son los recomendados en el mercado para la preparación de exámenes de la Universidad de Cambridge. Nueve de los participantes ya habían tomado años antes el examen llamado Preliminary English Test, PET, de la Universidad de Cambridge el cual es previo al First Certificate in English, FCE. Todos lo obtuvieron con un puntaje alto. El material utilizado en la enseñanza del inglés en la preparatoria, para la preparación del examen FCE, consiste en libros de texto, audio-casetes, diccionarios y ocasionalmente películas comerciales en inglés, pero no utilizan textos o materiales de TOEFL en absoluto, (véase tabla No. 8)

Tabla No. 8 CARACTERÍSTICAS DE LOS SUJETOS PARTICIPANTES						
Sujetos	Género	Promedio Edad	Años Estudio Inglés	Años en Escuelas Bilingües	Exámenes Cambridge	Estudios en el Extranjero
12	Fem	17.5	11.8	3 sujetos 8 años	5 sujetos PET	2 sujetos
11	Masc	17.8	12.9	5 sujetos 8.5 años	4 sujetos PET	4 sujetos
<b>Total Sujetos</b> 23		<b>Total Promedio</b> 17.6	<b>Total Promedio</b> 12.3	<b>Total Promedio</b> 8 / 8.2 años		

## 5.2 PROCEDIMIENTO

Una vez obtenida la autorización de la dirección de la escuela y la aceptación por parte de los sujetos para tomar las dos pruebas, TOEFL y FCE, durante su tiempo de clase, se procedió a explicarles en qué consistía el examen TOEFL y su método de aplicación para respetar tanto los tiempos estipulados y forma de contestar cada sección del mismo. Se revisó también la forma de aplicación de FCE para mantener todas las reglas a seguir en cuanto a lo que indica el manual de Cambridge ESOL Examinations, haciendo del conocimiento de los sujetos todos estos procedimientos, con el objeto de crear una atmósfera similar a la que se tiene cuando se aplican este tipo de exámenes. Previo a la aplicación de los exámenes se pidió a los sujetos que contestaran un cuestionario (ver apéndice 2) en el que se solicitaba información referente a su edad, estudios de inglés en el extranjero, en el país y/o en escuelas bilingües o con un sistema de enseñanza de inglés intensivo, así como si habían tomado algún examen de proficiencia de inglés. Se destinaron dos semanas para aplicar las dos pruebas, en la primera se aplicó FCE, una sección del examen por día y en la segunda semana el examen TOEFL, aplicando de igual manera una sección cada día. Al constar de tan solo tres partes este examen, comprensión auditiva, gramática y comprensión de lectura, se requirieron de tres días para esta prueba (ver apéndice 3).

Por razones de carácter personal, el examen de expresión oral del FCE solo accedieron a tomarlo diez sujetos. Dicho examen se aplicó en

pares, con dos maestros capacitados y autorizados por Cambridge para ser examinadores de expresión oral. La única sección que no se aplicó fue la de escritura ya que la calificación interjuez se lleva a cabo en Cambridge por profesores capacitados para tal fin, razón por la que asignar una calificación justa y válida constituía un problema al no contar ni con la escala de calificación ni la capacitación para evaluar esta sección.

### **5.3 CALIFICACIÓN DE LOS EXÁMENES**

La calificación de las pruebas se llevó a cabo siguiendo las instrucciones contenidas en las publicaciones de donde se tomaron los exámenes, tanto de TOEFL como de FCE, y de las Cambridge ESOL FCE Specifications (2007) con respecto al valor numérico que tienen los ítems en cada sección de las pruebas. Debido a que no se contó con los instrumentos utilizados por el ETS o por Cambridge ESOL Examinations para calificar los exámenes, se obtuvieron puntajes con características no oficiales pero que no obstruyeron el desarrollo de la investigación presentada en este trabajo. Sin embargo se pudieron obtener datos con características oficiales en la sección de Expresión Oral del FCE, ya que fue la única prueba que se calificó tanto con un instrumento oficial como por examinadores autorizados de Cambridge ESOL Examinations.

Se les mostraron los resultados obtenidos a los sujetos involucrados en el estudio y hubo coincidencia entre los resultados de sus exámenes y su desempeño escolar, ya que las secciones en las que reflejaron tener un bajo puntaje coincidía con el tipo de problemas que experimentaban en la materia de inglés con respecto a FCE. Su desempeño fue un tanto bajo en el examen TOEFL, nunca habían tomado este tipo de examen y las secciones de comprensión auditiva así como la de gramática fueron las que más dificultad representó para los sujetos.

## **6. RESULTADOS**

El objetivo de este estudio es comparar los puntajes obtenidos en el TOEFL y el FCE por un grupo de alumnos de preparatoria con el propósito de conocer su desempeño en ambas pruebas, ya que solo recibieron preparación para tomar el examen FCE.

Los resultados presentados en la tabla 9 muestran el comportamiento del grupo con respecto a cada subprueba tanto de TOEFL como de FCE. Los datos se expresan por media, desviación estándar, porcentajes de aciertos y rangos de mínimo a máximo. El total de ítems de cada subprueba se observa en la fila de cada prueba. En las dos pruebas receptivas del TOEFL, la de Comprensión Auditiva y la de Comprensión de Lectura, se encontraron ligeras similitudes en los porcentajes 68% y 60%, respectivamente. En cambio, en la subprueba de Escritura y Gramática los resultados están bajos, dado que sólo hay un 51% de aciertos.

<b>Tabla No. 9 PUNTAJES DE TOEFL Y FCE (N= 23)</b>					
<b>Examen</b>	<b>Media</b>	<b>%</b>	<b>Rango</b>		<b>Desviación Estándar</b>
			<b>Min.</b>	<b>Max.</b>	
<b>TOEFL Comp. Auditiva No. de Ítems: 50</b>	<b>34.04</b>	<b>68%</b>	<b>23</b>	<b>46</b>	<b>5.77</b>
<b>TOEFL Exp. Escrita y Gram. No. de Ítems: 40</b>	<b>20.56</b>	<b>51%</b>	<b>15</b>	<b>31</b>	<b>5.08</b>
<b>TOEFL Vocabulario Comp Lectura No. De Ítems: 60</b>	<b>36.00</b>	<b>60%</b>	<b>27</b>	<b>47</b>	<b>7.05</b>
<b>TOTAL TOEFL No. de Ítems : 150</b>	<b>90.60</b>	<b>60%</b>	<b>69</b>	<b>124</b>	<b>15.37</b>
<b>FCE Comprensión de Lectura No. de Ítems: 35</b>	<b>28.91</b>	<b>83%</b>	<b>21</b>	<b>35</b>	<b>3.84</b>
<b>FCE Gramática No. de Ítems: 65</b>	<b>42.34</b>	<b>65%</b>	<b>28</b>	<b>57</b>	<b>7.91</b>
<b>FCE Comp. Auditiva No. de Ítems: 30</b>	<b>20.56</b>	<b>68%</b>	<b>13</b>	<b>27</b>	<b>3.42</b>
<b>FCE Expresión Oral No. de Ítems (puntaje total) 25*</b>	<b>22.30</b>	<b>89%</b>	<b>19</b>	<b>25</b>	<b>2.41</b>
<b>TOTAL FCE No. de Ítems: 155</b>	<b>101.47</b>	<b>65%</b>	<b>68</b>	<b>141</b>	<b>20.19</b>

\* En esta subprueba no existe un número específico de ítems sino un puntaje total de 25 puntos.

En el FCE y con respecto a las pruebas de Comprensión Auditiva y de Lectura, se obtuvieron porcentajes más contrastantes que en el TOEFL en el mismo tipo de subpruebas, 68% y 83% respectivamente.

Adicionalmente, en Expresión Oral se obtuvieron tanto un porcentaje alto como un rango mínimo y máximo superior a todas las otras subpruebas de FCE. En lo referente a los rangos mínimos y máximos todas las subpruebas del TOEFL manifiestan un comportamiento muy similar, a diferencia de FCE donde los porcentajes más similares solo se encontraron en Gramática y Comprensión Auditiva. Estos resultados demuestran que un grupo de sujetos que fueron preparados para tomar el examen FCE y con un nivel de inglés intermedio alto fueron capaces de tener un desempeño con porcentajes altos en la prueba TOEFL.

## **7. 1 COMPARACION DENTRO PRUEBAS –**

### **CONFIABILIDAD**

Para conocer cómo se relacionan las variables entre pruebas en su conjunto y posteriormente conocer su consistencia interna por medio de la confiabilidad se llevó a cabo una correlación Pearson (véase tabla no. 10).

Tabla No. 10 Correlación de Puntajes Obtenidos en TOEFL y FCE (N= 23 FCE – Expresión Oral = 10)									
	Toefl Comp. auditiva	Toefl Exp. escrita gramática	Toefl Voc. comp. lectura	Total Toefl	Fce Comp. lectura	Fce Gramática	Fce Comp. auditiva	Fce Exp. Oral	Total Fce
Toefl comprensión Auditiva	1	.513*	.583**	.813**	.530*	.602**	.470*	.260	.700**
toefl expresión escrita y Gramática		1	.693**	.842**	.425	.663**	.299	.743*	.480*
toefl vocabulario comprensión de lectura			1	.907**	.429	.617**	.358	.863**	.495*
Total Toefl				1	.537*	.729**	.440*	.823 **	.653**
Fce comprensión de lectura					1	.635**	.418*	.429	.640**
Fce gramática						1	.576**	.673*	.809**
Fce comprensión Auditiva							1	.099	.653**
Fce expresión oral								1	.664*
Total Fce									1

\* La Correlación es significativa al nivel de 0.05 level (2-colas).

\*\* La Correlación es significativa al nivel de 0.01 level (2-colas)

Considerando el aspecto de confiabilidad en referencia a consistencia interna, la prueba TOEFL mostró esta característica ya que sus subpruebas operan de manera similar según se puede ver en los datos,  $r=.513$ ,  $r=.583$ ,  $r=.693$ ,  $r=.842$ . Este aspecto coincide con FCE en dos de sus subpruebas, Comprensión de Lectura  $r=.635$  y Gramática  $r=.576$ , pero en el caso de Expresión Oral, los datos no muestran una consistencia interna fuerte, aunque si bien en esta última subprueba sólo participaron diez sujetos y probablemente eso sea el origen del resultado que arroja el estadístico respecto a la comparación entre pruebas.

Esta situación de inconsistencia presente en FCE no se presentó en ninguna de los componentes de TOEFL concluyendo así que TOEFL demostró tener una consistencia interna fuerte.

Posteriormente se llevó a cabo una comparación entre ambas pruebas para conocer como se correlacionaban entre si.

## **6.2 Comparación entre Pruebas - Validez Concurrente**

La validez concurrente de las pruebas se obtuvo al contrastar cada subprueba del TOEFL con las correspondientes del FCE y debido a que estas dos pruebas de proficiencia de lengua ofrecen escalas de medición referentes a las habilidades evaluadas, esto permitió llevar a cabo la correlación de cada una de sus subpruebas y conocer su validez.

Al revisar los totales de la correlación se encontró que TOEFL y FCE de manera general tienen una fuerte correlación en la suma total de sus subpruebas,  $r=.729$ ,  $r=.823$  y  $r=.653$ , con una significancia alta de  $p < 0.01$ . Pero tuvo  $r=.537$  y  $r=.440$  con una significancia media de  $p < 0.05$  en dos de los totales de las subpruebas. En cuanto a cada una de las subpruebas de TOEFL en su correlación con FCE, en la subprueba de Comprensión Auditiva TOEFL se obtuvo  $r=.530$  y  $r=.470$  con Comprensión de Lectura y Comprensión Auditiva de FCE con una significancia media de  $p < 0.05$  respectivamente. Pero Comprensión Auditiva TOEFL obtuvo  $r=.602$  en correlación con Gramática FCE y en el total FCE  $r=.700$ , ambas con una significancia alta de  $p < 0.01$ .

Una de las subpruebas que se caracteriza por reflejar el uso y la producción de la lengua como es la Expresión Escrita y Gramática de TOEFL tuvo una baja correlación con dos pruebas que examinan habilidades que tienen en común que son principalmente receptivas, Comprensión de Lectura de FCE que obtuvo  $r=.425$  y Comprensión Auditiva de FCE con  $r=.299$  ambas sin factor de significancia.

Pero Expresión Escrita y Gramática de TOEFL en correlación con Expresión Oral de FCE se obtuvo  $r=.743$  y  $r=.480$  en el total de FCE con una significancia media de  $p < 0.05$ . Por último en la correlación de Vocabulario y Comprensión de Lectura TOEFL con Comprensión de Lectura FCE se obtuvo  $r=.429$  y  $r=.358$  en la Comprensión Auditiva de FCE pero no hubo datos de significancia en este caso.

Sin embargo, la subprueba de Vocabulario y Comprensión de Lectura TOEFL obtuvo  $r=.617$  en correlación con Gramática FCE así como  $r = .863$  en Expresión Oral de FCE con una alta significancia de  $p < 0.01$ . En la correlación de Vocabulario y Comprensión de Lectura y el total de FCE se obtuvo  $r = .495$  con una significancia media de  $< p 0.05$ .

Se encontró evidencia, al hacer la comparación entre pruebas, de que existe una fuerte correlación entre varias de las subpruebas, principalmente en las tres subpruebas de TOEFL con la de Gramática de FCE, donde todas ellas muestran una alta significancia. Aunque en la correlación de las otras subpruebas de ambos instrumentos se obtienen valores que varían en su significancia, también existen elementos de

coincidencia y con una significancia que va desde un nivel medio hasta uno alto y que muy probablemente sea a consecuencia de la semejanza en los constructos de cada subprueba. En las instancias en las que se tiene una alta o mediana significancia en la correlación, se podría decir que esta relación tiene que ver con la propuesta de Bachman (1990) quien afirma que para que los valores obtenidos sean válidos en una correlación, éstos deben ser indicativos de que se hizo una comparación entre pruebas que examinan el mismo tipo de habilidades.

## **7. CONCLUSIONES**

La confiabilidad, como se estableció en el apartado 2.7, es un elemento importantísimo en las pruebas ya que puede haber confiabilidad sin validez, pero nunca este último aspecto sin confiabilidad por lo que es necesario retomar la noción de confiabilidad como la consistencia de las puntuaciones obtenidas por los mismos sujetos, cuando son examinados con el mismo instrumento de medición o con conjuntos equivalentes de reactivos, Anastasi y Urbina (1998).

En la introducción de esta investigación, se plantean tres hipótesis, en la primera de ellas se declara que se corroborará que la consistencia interna del FCE es menos fuerte en cuatro de sus subpruebas con respecto a tres del TOEFL.

Al llevar a cabo la comparación entre estas dos pruebas, se encontró que efectivamente la prueba TOEFL obtuvo datos que revelan que tiene una fuerte consistencia interna en sus subpruebas, donde el coeficiente de correlación ( $r$ ) tuvo una significancia con tendencia alta en todas las pruebas. Únicamente en el caso de Expresión Escrita y Gramática en su correlación con Comprensión Auditiva, tiene una significancia menos alta,  $p < 0.05$  lo que significa que la consistencia es baja.

Ahora bien, en el caso de FCE, los datos indican que este examen tiene una consistencia interna fuerte con un coeficiente de correlación con una alta significancia, en los totales de las subpruebas. Las subpruebas más fuertes son Gramática y Comprensión de lectura,  $p < 0.01$ , pero con una correlación baja con las demás subpruebas. Estos datos coinciden con lo reportado por Bachman et al (1988) corroborando la primera hipótesis, en la que se plantea que TOEFL tiene una consistencia interna más fuerte en las tres subpruebas comparadas. En su análisis informan los investigadores de referencia, que hubo una consistencia fuerte y equivalente entre dos versiones de la subprueba de Comprensión de

Lectura aplicadas en su estudio, lo cual coincide con los datos del presente trabajo.

Pero no fue así en el caso de las dos subpruebas de Comprensión Auditiva que utilizaron, ya que encontraron que tenían baja consistencia interna y falta de equivalencia entre esas dos diferentes ediciones de las subpruebas. En el caso del presente trabajo también se encontró que Comprensión Auditiva tuvo una consistencia interna baja. Bachman et al (1988) atribuyeron lo anterior en su caso, a problemas causados por el sistema de calificación de la prueba misma, ya que solo se agregan los puntajes de esta subprueba al total general. En cuanto a la consistencia interna de las subpruebas de Escritura y de Expresión Oral de FCE, encontraron que no les era posible estimar el grado de consistencia interna de estas dos secciones debido a su amplio rango de respuestas. En el caso de este trabajo también se obtuvo una consistencia interna baja en Expresión Oral.

La explicación que se propone para estos resultados, es que el FCE es una prueba con una orientación comunicativa, razón por la que su diseño ofrece contextos tanto de interacción como situacionales, más auténticos, es decir más apegados a la vida real. Este aspecto tiene como resultado que los componentes de la prueba midan una amplia variedad de habilidades relativas a la proficiencia de la lengua, pero a cambio de esta gama controlada de posibilidades de producción de la lengua, no se tiene una alta ínter correlación de los componentes de este

instrumento. El conjunto de rasgos implicados en la construcción de la prueba FCE, tiene como consecuencia también, que el tiempo requerido para resolver cada subprueba dependerá de su objetivo y de los ítems que contenga.

Consecuentemente, la confiabilidad de la prueba FCE en el sentido de consistencia interna, puede presentar problemas al tratarse de un instrumento que contiene variedad y heterogeneidad. Wood (1993:138) argumenta al respecto que la consistencia interna de una prueba se ve comprometida, cuando de manera deliberada se tiene una prueba con una mezcla de modalidades, habilidades y contenidos, pero que a cambio dicho examen permite que el candidato tenga mejores resultado al tener más oportunidades para demostrar sus habilidades lingüísticas. Por lo tanto cuando se habla de confiabilidad en FCE, los investigadores de validez de la Universidad de Cambridge dicen que se trata de una confiabilidad compuesta ya que esta noción abarca varios componentes y su inter correlación, la cual entre más fuerte reflejará una confiabilidad compuesta más fuerte. Agregan que, cuando se habla de consistencia interna en FCE, los estimados para este fin dependen de la homogeneidad de los ítems que contiene cada subprueba, además de que las diversas tareas y amplio espectro de los enfoques que incluyen las pruebas, representan dificultad para obtener alphas altos en estudios de confiabilidad de FCE, aunque aclaran que en los diferentes estudios que

han llevado a cabo han encontrado estimados de confiabilidad altos y consistentes, de 0.94. (Research Notes. Issue 15 / February 2004:21-23).

Finalmente es necesario agregar que uno de los aspectos de mayor relevancia para la Universidad de Cambridge es la confiabilidad de aplicación de sus pruebas. Para ello se tiene un programa permanente de revisión y actualización de los lineamientos de aplicación de las pruebas, del personal y los centros que las aplican, del sistema de recolección y el envío de las pruebas a la Universidad de Cambridge para su calificación.

Por todo lo anterior se acepta la hipótesis número uno.

En cuanto a la segunda hipótesis de este trabajo, donde se afirma que el FCE tiene una validez con una alta significancia a diferencia del TOEFL, antes de proceder a exponer las conclusiones a este respecto, es necesario retomar el concepto de validez concurrente el cual es el grado en que los puntajes de una prueba están relacionados con los puntajes de otro instrumento de medición que se administra al mismo tiempo y a los mismo sujetos, Bachman (1990).

En la correlación que se llevó a cabo entre TOEFL y FCE, se puede apreciar que el TOEFL, en dos de sus pruebas y Gramática del FCE, tienen una validez significativa lo cual contradice la segunda hipótesis de este trabajo, en la que se dice que el TOEFL no tiene, en su totalidad, una validez con una alta significancia ante el FCE.

Ambas pruebas miden la capacidad de comprensión de información hablada y escrita, pero existen diferencias importantes en cuanto a los contenidos entre una y otra prueba así como en sus respectivos formatos. Esta situación también fue estudiada por Bachman et al (1988) en su estudio de comparabilidad entre estas dos pruebas. La Comprensión Auditiva en el TOEFL tiene un objetivo muy simple y directo, medir la habilidad del candidato para entender la lengua inglesa tal y como se habla en los Estados Unidos de Norteamérica, el cuestionario evalúa la comprensión de las ideas principales en el texto y el orden de las ideas que apoyan las respuestas seleccionadas. El formato para contestar el examen utilizado como siempre es el de opción múltiple. En lo que respecta a FCE, el objetivo de este tipo de subprueba es medir la habilidad de comprensión de la lengua hablada con diferentes tipos de acentos del Reino Unido, edades y contextos. Por lo tanto el candidato resuelve diferentes tareas que incluyen varios aspectos: identificar información específica, hacer inferencias, comprender la información más importante en diálogos, monólogos, conferencias, etc., que han sido tomados de materiales auténticos tales como programas de radio, conferencias, anuncios, entre otros. El FCE, a diferencia del TOEFL, tiene un formato en el que incluye opción múltiple, tomar notas, completar espacios en blanco con una palabra o una frase, combinación de opciones y selección de respuestas entre dos y tres opciones. Finalmente el tiempo de aplicación es similar en ambas pruebas, 40 minutos.

Dentro de las recomendaciones que presentan Bachman et al (1988) en su estudio, hacen mención de que una prueba sicométrica de las características del TOEFL es más fácil de calificar y consistente en sus resultados. El FCE como se ha dicho, por tener un enfoque comunicativo puede originar el tipo de resultados que se reportaron en su investigación, ya que el candidato puede producir más de una respuesta en algunos casos. Además al tener un espectro de respuestas más amplio en algunas secciones de la prueba, se pueden presentar diferencias en las variables al hacer una correlación de FCE con algún otro instrumento de medición.

Pero una subprueba de FCE, Gramática, tuvo un coeficiente de correlación  $r$  con una significancia alta con todas las subpruebas de TOEFL a pesar de las diferencias en cuanto al tipo de habilidad que cada subprueba involucra, por ejemplo la correlación entre la Comprensión Auditiva de TOEFL. Una probable explicación en este caso sería que la subprueba de Gramática de FCE tiene cinco secciones que implican un conocimiento de estructuras gramaticales en diferentes contextos, donde el candidato debe identificar errores gramaticales, uso de vocabulario en un contexto, completar textos o bien parafrasear enunciados entre otras tareas. Por lo anterior y siendo el conocimiento de la gramática y de vocabulario aspectos fundamentales para comprender información dentro de un contexto auditivo, se podría afirmar que esas características de la

subprueba de Gramática de FCE propiciaron una correlación con una significancia alta también con las otras dos subpruebas de TOEFL, donde existen elementos coincidentes en cuanto a la naturaleza de esas subpruebas y la de FCE en cuestión.

Estas dos pruebas se correlacionaron y revelaron tener una validez concurrente con una significancia alta en varias de sus subpruebas. Pero se rechaza la segunda hipótesis al no existir datos que permitan afirmar que el FCE tiene una validez con una alta significancia, en las cuatro subpruebas que se estudiaron, ante las tres del TOEFL.

En la tercera hipótesis se asevera que los candidatos tendrán un mejor desempeño en FCE ya que se han preparado para tomarlo. Los sujetos resolvieron la prueba TOEFL sin haber tenido experiencia con este tipo específico de examen.

La hipótesis se acepta ya que los resultados obtenidos en el FCE fueron más altos en dos subpruebas y en el total de puntajes, que los logrados en el TOEFL. La preparación para el FCE involucró conocimiento de la lengua para realizar tareas comunicativas, cotidianas, con tareas que involucran materiales auténticos, es decir tomados de periódicos, revistas, folletos, entre otros. Mientras que en el caso de TOEFL el material utilizado en esa prueba, es el comúnmente usado en la vida cotidiana, pero en un contexto académico universitario. Sin embargo es pertinente

aclarar que esta aseveración no implica que ambas pruebas sean de igual dificultad ya que no se tiene evidencia que pueda confirmarlo. Pero a juzgar por los resultados en los puntajes se podría concluir que los sujetos desarrollaron su proficiencia en la lengua inglesa y fueron capaces de resolver un examen con un diseño, formato y metodología de administración completamente diferente al FCE, hecho que confirma lo expuesto en el capítulo 3.1 referente a la proficiencia en una segunda lengua.

## BIBLIOGRAFÍA

- American Psychological Association. (1985). American Educational Research Association and National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC: American Psychological Association
- Aiken, R. L. y Lewis, T. (1996). Tests Psicológicos y Evaluación. 8ª Edición. Prentice Hall, México.
- Alderson, J.C., Clapham, C. y Wall, D. (1995). Language Test Construction and Evaluation. Cambridge Language Teaching Library. Cambridge: Cambridge University Press.
- Anastasi, A. y Urbina, S. (1998). Tests Psicológicos. 7ª Edición. Prentice Hall. México.
- Association of Language Testers in Europe. (1994). The ALTE Code of Practice. Cambridge: Cambridge University Press.
- Association of Language Testers of Europe (ALTE, 2004). Disponible en: [www.alte@ucles.org.uk](mailto:www.alte@ucles.org.uk)
- Bachman, L.F., Kunnan, J. A. y Lynch, B. (1988). Ability and Task Analysis as a Basis for Examining Content and Construct Comparability in Two Proficiency Test Batteries. 10<sup>th</sup> Annual Language Testing Research Colloquium. University of Illinois at Urbana – Champaign. March, 5 to 7, 1988.
- Bachman, L. F. (1990) Fundamental Considerations in Language Testing. Oxford, UK: Oxford University Press.

- Bachman, L.F., Davidson, F., Ryan, K. y Choi, I-C. (1995). An Investigation in the Comparability of two Tests of English as a Foreign Language: The Cambridge - TOEFL Comparability Study. Cambridge: Cambridge University Press.
- Bachman, F. L., y Palmer, A. (1996) Language Testing in Practice. Oxford, UK: Oxford University Press.
- Baker, C. (1997). Foundations of Bilingual Education and Bilingualism. 3ra Edición Bilingual Education and Bilingualism 27. Series Eds. Nancy Hornberger y Colin Baker. Multilingual Matters, Ltd. USA
- Butler, G.Y., y Hakuta, K. (2004). The Handbook of Bilingualism. Bilingualism and second language acquisition. (Eds.). T. K. Bathia y W. C. Ritchie. Blackwell Publishing, Ltd. Australia.
- Cohen, J.R. y Swerdlik, E.M. (1996) Pruebas y Evaluación Psicológicas. 4ª Edición. McGraw Hill. México
- Cohen, J.R., Swerdlik, E.M. (2000) Pruebas y Evaluación Psicológicas. Introducción a las Pruebas y a la Medición. 4ª Edición. McGraw Hill: México.
- Cambridge ESOL. First Certificate Exam Specifications: Disponible en: [www.cambridgeesol.com](http://www.cambridgeesol.com)
- Campbell, D.T. y Fiske, D.W. (1959). Convergent and Discriminant Validation by The Multitrait-Multimethod Matrix. Psychological Bulletin, LVI (1959) Educational and Psychological Measurement. Science Research Associates, American College Personnel Association.
- Canadian Academic English Language Assessment (CAEL): Disponible en: [www.assessment.cael@carleton.ca](mailto:www.assessment.cael@carleton.ca)

- Clark, C. (2002). The Common European Framework of Reference. Centro Linguistico di Ateneo. Universita degli Studi di Padova
- Council of Europe. (2003a). Common European Framework of Reference for Languages: Learning Assessment, Cambridge: Cambridge University Press
- Cummins, J. (1980). Communicative Competence Approaches to Language Proficiency Assessment: Research and Application. Linguistic Interdependence Among Japanese and Vietnamese Immigrant Students. (Eds.). Charlene Rivera . Multilingual Matters Ltd. 9. USA
- Cummins, J. (1980). An Introductory Reader to the Writings of Jim Cummins. Ed. Colin Baker y Nancy H. Hornberger. Bilingual Education and Bilingualism 29. Multilingual Matter Ltd. USA
- Cummins, J. (2000). Language, Power and Pedagogy. Bilingual Children in The Crossfire. Multilingual Matters Ltd. Antony Rowe Ltd., Chippenham, Wiltshire. Great Britain.
- Del Vecchio, A. y Guerrero, M. (1995). Handbook of English Language Proficiency Tests. EAC-West, New Mexico Highlands University, Albuquerque. Educational Publications. December 1995.
- Des Brisay, M. (1989). Two Tests are Better than One? Eleventh Annual Language Testing Research Colloquium. Defense Language Institute English Language Center. San Antonio, Texas.
- Educational Testing Service, ETS. (2006). Research Agenda. Assessment and Testing. Cambridge: Cambridge University Press

ETS Development Research Agenda. (2006). Test of English as a Foreign Language. Princeton, NJ: Educational Testing Service  
ETS: Mapping Test Scores onto the Common European Framework.  
Disponible en: [www.ets.org/vgn-ext-templating](http://www.ets.org/vgn-ext-templating)

Educational Testing Service. TOEFL. Disponible en: [www.ets.org/toefl](http://www.ets.org/toefl)

First Certificate in English Handbook. (2001). University of Cambridge. Local Examinations Syndicate, (UCLES). Cambridge ESOL Examinations. Cambridge. UK

First Certificate in English Handbook for teachers. . (2005 ). University of Cambridge ESOL Examinations. (UCLES 2005). Cambridge. UK

Magnan, S.S. (1986). Assessing Speaking Proficiency in the Undergraduate Curriculum. Data from French. Foreign Language Annals, 10, 429-438.

McNamara, T. (2000). Language Testing. Oxford University Press: Oxford.

Messick, S.A. (1989). Educational Measurement, News York: MacMillan.

Messick, S.A. (1996b). Validity of Performance Assessment. In Philip, G. (1996). Technical Issues in Large-Scale Performance Assessment. Washington. DC: National Center for Educational Statistics.

Michigan English Language Assessment Battery, (MELAB) The MELAB:  
Disponible en: [www.elicertifumich.edu](http://www.elicertifumich.edu)

Nunan, D. (1995) Action Research in the Language Classroom Research. Cambridge: Cambridge University Press.

Oller, J. W., y Perkins, K. (1980). Research in Language Testing. Rowley, MA: Newbury House.

Red Iberoamericana para la Acreditación de la Calidad de la Educación Superior. Glosario. Disponible en: [www.riacesnet/glosarioc.html](http://www.riacesnet/glosarioc.html)

Salvia, J. y Yssel dyke. J. (1995) Evaluación en la Educación Especial. México. Manual Moderno (1997 español).

Shrum, L. J., y Glisan, W.E. (2005). Teacher's Handbook. Contextualized Language Instruction. Third Edition: Thomson Heinle.USA.

The IELTS test structure. (2005). Disponible en: <http://www.ielts.org>

TOEFL Scoring Guide (2005) Princeton, NJ: Educational Testing Service. Disponible en: <http://www.ets.org/portal/site/ets/menuitem>

University of Cambridge. ESOL Examinations. (2005) Research Notes: Issue 22, 2-5.

Wood, R. (1993). Assessment and Testing. Cambridge: Cambridge University Press.

**Learners & Test Takers Paper-Based Testing (PBT)**

**TOEFL SCORING GUIDE**

**TOEFL PBT Scores**

<b>Skill</b>	<b>Score Range</b>
Listening Comprehension	31 to 68
Structure/Written Expression	31 to 68
Reading Comprehension	31 to 67
Total Score	310 to 677

The raw score for each section is converted by statistical means to a number on what is called the TOEFL test scale. The TOTAL paper-based test score is reported on a scale that ranges from 310 to 677.

**TWE® (Test of Written English™) Scores**

The TWE score is reported on a scale of 1 to 6 and printed in a separate box on the TOEFL score report. It is not added to the TOEFL score. A score between 2 points on the scale (for example, 5.5, 3.5) can also be reported

**Computer- and Paper-Based Test Scoring Guide**

The following are guidelines for scores relating to the Writing portion of the TOEFL® CBT/PBT test:

**Score of 6**

An essay at this level

- effectively addresses the writing task
- is well-organized and well-developed
- uses clear, appropriate details to support a thesis or illustrate ideas
- displays consistent facility in the use of language
- demonstrates syntactic variety and appropriate word choice.

**Score of 5**

An essay at this level

- may address some parts of the task more effectively than others
- is generally well-organized and developed
- uses details to support a thesis or illustrate an idea
- displays facility in the use of the language
- demonstrates some syntactic variety and range of vocabulary.

**Score of 4**

An essay at this level

- addresses the writing topic adequately but may slight parts of the task
- is adequately organized and developed
- uses some details to support a thesis or illustrate an idea

demonstrates adequate but possibly inconsistent facility with syntax, and usage may contain some errors that occasionally obscure meaning.

**Score of 3**

An essay at this level may reveal one or more of the following weaknesses:

- inadequate organization or development
- inappropriate or insufficient details to support or illustrate generalizations
- a noticeably inappropriate choice of words or word forms
- an accumulation of errors in sentence structure and/or usage.

**Score of 2**

An essay at this level is seriously flawed by one or more of the following weaknesses:

- serious disorganization or underdevelopment
- little or no detail, or irrelevant specifics
- serious and frequent errors in sentence structure or usage
- serious problems with focus.

**Score of 1**

An essay at this level

- may be incoherent
- may be undeveloped
- may contain severe and persistent writing errors.

**Score of 0**

An essay will be rated 0 if it

- contains no response,
- merely copies the topic,
- is off-topic, is written in a foreign language, or consists only of keystroke characters.

**TWE® Scores**

The TWE score is reported on a scale of 1 to 6. It is printed in a separate box on the TOEFL score report. It is not added to the TOEFL score.

ETS: <http://www.ets.org/portal/site/ets/menuitem>

TOEFL Scoring Guide 2005

## APÉNDICE 2

### BASE DE DATOS DE ALUMNOS QUE TOMARON LOS EXÁMENES FCE Y TOEFL

NOMBRE \_\_\_\_\_

EDAD \_\_\_\_\_

SEXO \_\_\_\_\_

AÑOS DE ESTUDIO DE INGLÉS \_\_\_\_\_

¿DONDE HAS ESTUDIADO INGLÉS?

¿HAS TOMADO ALGÚN EXAMEN DE LA UNIVERSIDAD DE CAMBRIDGE?

EN CASO AFIRMATIVO, MENCIONA CUÁL EXAMEN TOMASTE Y CUÁNDO

¿HAS ESTUDIADO EN ESCUELAS BILINGÜES?  
EN CASO AFIRMATIVO, MENCIONA CUALES Y CUANTOS AÑOS

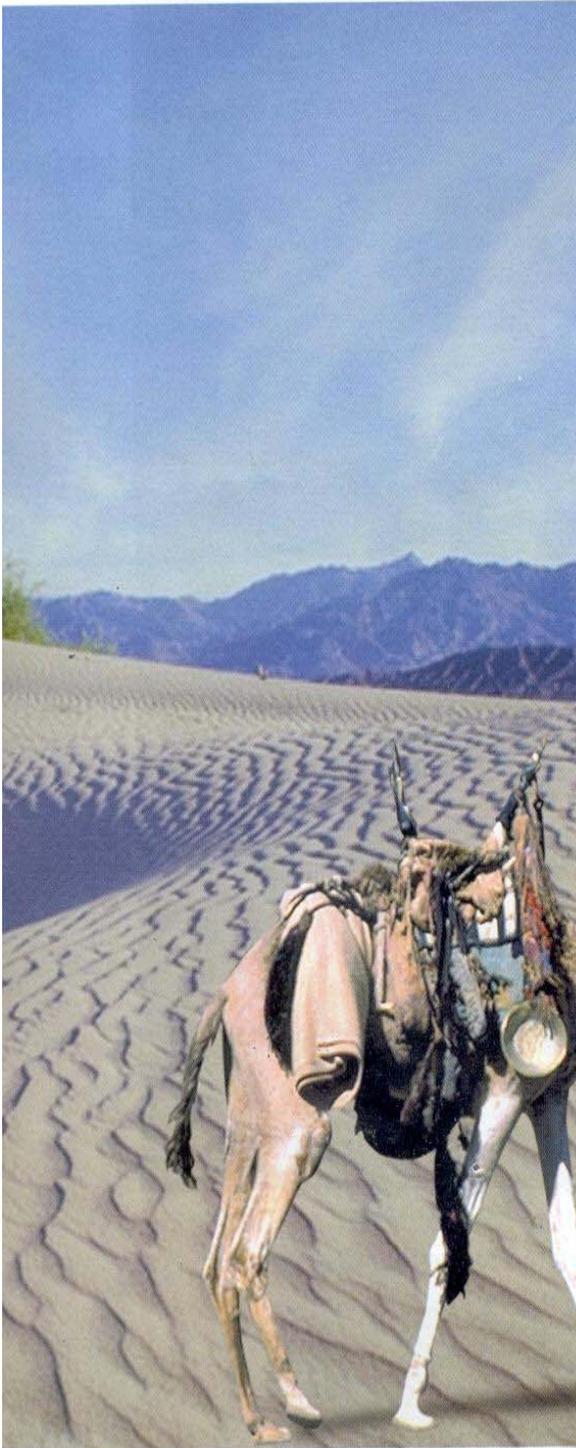
¿HAS ESTUDIADO EN EL EXTRANJERO?  
EN CASO AFIRMATIVO, MENCIONA DÓNDE, CUÁNTO TIEMPO Y QUÉ ESTUDIASTE

Gracias por tu colaboración.

## PAPER 1 • READING

## Part 1

You are going to read a newspaper article about a desert in Israel. Choose the most suitable heading from the list (A-I) for each part (1-7) of the article. There is one extra heading which you do not need to use. There is an example at the beginning (0).



**A** Viewing the desert as a threat.

**B** Illusions and marvels.

**C** Travelling into the wilderness.

**D** Dangers beneath the desert.

**E** The many faces of the desert.

**F** A changing way of life.

**G** Land of the Bible.

**H** Bringing life to the desert.

**I** An undiscovered land.

# Desert Heat



0

Only 80 kilometres to the south of Jerusalem, the Negev desert begins. It covers nearly half of Israel, yet it is largely ignored by the country's visitors, who see it as nothing but endless dust and rock. Traditionally, tourism has only affected the outskirts of the desert : the Red Sea in the south and the Dead Sea in the north.

1

With temperatures reaching 45°C, this dried-up landscape of dwarfed bushes and threatening mountains will always terrify some travellers. In the total silence it is common for people to believe they hear dogs barking, phones ringing and buses stopping. But these hallucinations soon wear off, and many people who visit the desert discover its wonders.

2

The Negev is a predominantly rocky desert, with an amazing variety of landscapes: flat, stony plains, canyons, plantless mountains and salt marshes. Steep cliffs rise above

dry riverbeds. Here and there are deep holes, the remains of copper mines made by the Egyptians 6,000 years ago. The overall impression is of a prehistoric landscape.

3

Nowadays, camels, the traditional desert transport, are being replaced by the less stubborn four-wheel – drive vehicle. Desert safaris, whether by car or camel, are now a growing industry. Treks range from an hour on a camel to seven-night journeys on foot or by jeep, with a mattress at night under the stars. How far you travel into the Negev and away from civilisation is your decision.

4

Despite its potential for tourist development, Israel has always been a little worried by the desert. David Ben-Gurion, Israel's first prime minister, said, "If the State does not put an end to the desert, the desert may put an end to the State."

5

Ben-Gurion's dream was to turn the desert into agricultural land, using a

system of underground water points. Kilometre by kilometre, the green squares of land, worked by Kibbutzim (co-operative farms) are turning dry rock into plantations of fruit. But the survival of the project is heavily dependent on the annual rainfall. Water has always been the most important survival factor for all life in the Negev.

6

A large proportion of the desert is also used by the Israeli army for training purposes, and it is rumoured that somewhere below the desert, Israel's nuclear arms station is hidden. The Negev is a fragile area which can be damaged easily. Rubbish left behind by the army is a common sight.

7

The tents of the Bedouins - the Arab people of the desert - can still be seen in some places. Increasingly, however, they are giving up their wandering lifestyle to live in permanent housing. But the herds of sheep and goats that provide the Bedouins with a source of income are still to be seen amongst the rocks of the desert.

## Part 2

You are going to read an article about H. G. Wells. For questions 8-15, choose the answer (A, B, C or D) which you think fits best according to the text.

H. G. Wells was born in 1866 in Bromley. He claimed to have a very ordinary brain, but in fact he predicted air attacks and atomic bombs long before they existed. He took no pleasure in being right though. Instead it just added to his growing feeling of pessimism. His last book was *Mind at the End of its Tether*, a work full of despair. In his final year he spent his time painting a mural in his home which showed the process of evolution, ending with the figure of Man. Beneath this figure he wrote "Time to Go".

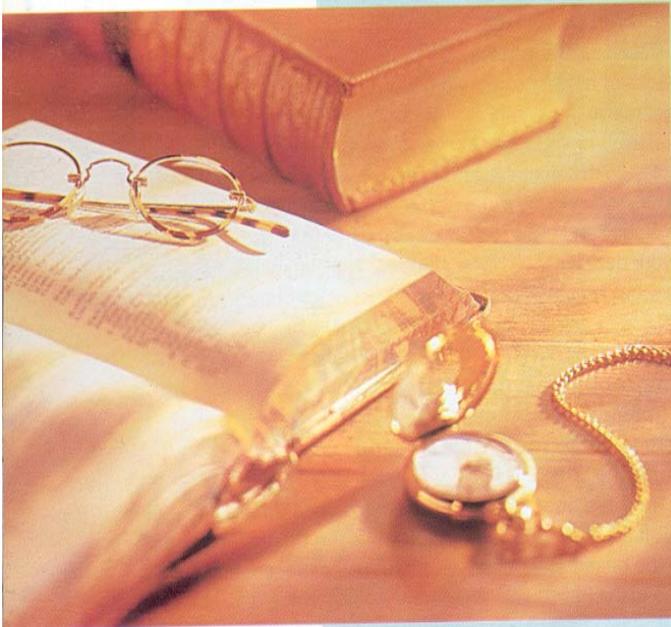
Wells' parents were, at various times, shopkeepers and servants. Wells was destined to become a shop assistant, but after suffering an accident as a child he was forced to spend a long time in bed. This period of inactivity gave him a love of reading which developed his imagination. After leaving school he first worked in a curtain shop, then trained as a teacher and biologist, but after this decided to write books for a living. At the age of 27, working by candlelight in a room in Kent (his landlady complained he used too many candles), he wrote the book that made his name, *The Time Machine*. It is the story of an unnamed time traveller who meets strange people in the future and witnesses the end of the world. It was a great success with Victorian readers.

Wells went on to produce "scientific romances" and short stories which were serialised in publications like the *Strand Magazine*. Many science fiction themes - aliens coming to earth, planetary disasters and so on - were dealt with in these early stories. Wells' ability to create such original work was amazing. He turned everyday events into incredible fantasies: a conversation about colonialism became *The War of the Worlds* and a walk round London was turned into *The Invisible Man*. Some of the predictions made in his books included the use of aeroplanes and tanks in war, the rise of the middle class, the liberation of women and the need for a world state.

His own background gave him the basic materials for his best novels, which were realistic comedies of lower-middle-class life. In these he was at his peak as an artist. Though he continued to write novels, his talent as a fiction writer was gradually overwhelmed by his enthusiasm for scientific knowledge and social change. His idea was that civilisation was "a race between education and catastrophe", and though

the catastrophe of the First World War damaged his optimism, he continued to fight for social change.

Wells' most important act is probably his least known. At the start of World War II, Wells produced a statement of human rights and formed a committee to work on this. The Declaration of the Rights of Man was published in many different languages and stimulated discussion worldwide. After the War, this document greatly influenced the wording of the United Nations Declaration of Human Rights.



**8** What made Wells become displeased in later life?

- A having an ordinary brain
- B making wrong predictions
- C being right about the future
- D being a pessimist

**9** What type of work was Wells originally destined to do?

- A administration
- B writing
- C cleaning
- D sales

**10** What helped his imagination to grow?

- A reading
- B listening to his parents
- C making predictions
- D writing books

**11** Wells' first book

- A was based on his own experience.
- B was written in London.
- C did not interest readers.
- D gained him popularity.

**12** What sort of stories did Wells write at first?

- A romantic
- B futuristic
- C war
- D mystery

**13** Wells' "incredible fantasies" were developed from

- A alien invasions of earth.
- B exploring planets.
- C ordinary incidents.
- D scientific predictions.

**14** What does "these" in line 35 refer to?

- A novels
- B basic materials
- C social classes
- D backgrounds

**15** How did the First World War affect Wells?

- A It changed his writing style.
- B He became more hopeful.
- C He stopped writing novels.
- D He became less positive.



## Part 3

You are going to read an article about Disneyland. Seven paragraphs have been removed from the article. Choose from the paragraphs (A-H) the one which fits each gap (16-21). There is an example at the beginning (0).

### Disneyland Paris

Think of Walt Disney. Think of his films. Think of the pure fantasy he created. Imagine entering the world of fantasy and perhaps even living part of it. In 1955, that's what Disney did for the world. He created a fantastic dream world that everybody could be part of: Disneyland.

0

D

Disneyland was magical and it was here to stay. The mind that envisioned this utopian ideal was indeed unique, and even after Disney's death the ideal continued. 1971 saw the opening of Disneyworld in Florida, and in 1983 Tokyo Disneyland opened in Japan. These worlds of fantasy attract millions of visitors every year, especially from Europe.

16

In 1988, in an enormous area just outside Paris, workmen broke ground and set about creating a new and incredible world of fantasy: EuroDisney. It was to be the most impressive creation ever. Not only would it contain all the essential elements that make up Walt Disney's ideal world of fantasy, but it would cater to people's everyday needs as well.

17

The person the Disney chiefs finally settled on to head the creation of EuroDisney seemed at first to be something of a strange choice. Bob Fitzpatrick did not come from a strictly business background, as you would expect of someone running a billion dollar enterprise. In fact, quite the opposite. Before he came to EuroDisney, Fitzpatrick was President of the California Institute of Arts.

18

He took a degree in medieval French and spent

some years teaching the language and eventually married a French woman. But should it really be a surprise that someone with an arts background should be chosen to build something like EuroDisney? Not when you consider that this is a world of beauty and fantasy.

19

This is the place where dreams can come true and the castle is the ultimate fairy tale. This is what Bob Fitzpatrick was responsible for creating. His main objective was not to build a park based on profit, but on the emotions the creations can produce in visitors.

20

The music they play is designed to delight children and take adults back to their childhood. People recognise the tunes and happy memories are triggered. Fitzpatrick believes that running EuroDisney is like running a huge theatrical production.

21

But although Fitzpatrick has created this incredible world of fantasy, the business side of things has not been neglected. Huge amounts of gifts and souvenirs are sold. There are thirty five restaurants and one of Europe's largest hotel operations.

Fitzpatrick even put his art background to good use in the hotels. At one particular hotel he tried 27 different shades of pink for the curtains before he was satisfied. He wanted everything to be just so.

This is surely a measure of the man's talent for creating something quite unique. Apart from the sheer beauty and excitement of EuroDisney, nothing was too trivial for Bob Fitzpatrick to deal with in his quest for perfection.

- A** So the bosses of Disney decided to create another world of fantasy, this time in Europe. They chose France as the location of the new site, and work began.
- B** Atmosphere is everything. Fitzpatrick believes that as people walk through the gates reality should be suspended, and it is not only through the eyes that visitors are drawn into the enchantment. The sense of hearing also plays an important role.
- C** But his roots were even further from big business. He spent seven years training to be a Jesuit priest. Then he decided that his real vocation was in teaching.
- D** When it first opened, the public went wild about it. Many thought its popularity was a passing phase, but several decades have proved that it wasn't.
- E** It should also come as no surprise that the design of Sleeping Beauty's Castle is taken from a medieval French manuscript. The castle is just part of EuroDisney's centrepiece, Fantasyland.
- F** Of course EuroDisney is much smaller than the American versions, but there is just as much variety. Fitzpatrick says it was designed with European tastes and habits in mind. "In Europe, everything is smaller," he says.
- G** Included in the site were hotels, restaurants, shops, campsites and transport facilities. The area was one fifth of the size of Paris and it would take nothing less than a genius to run it.
- H** The staff are like a cast of 170,000 actors. They are all young, and they obviously enjoy the roles they play. It is pure show business.



## Part 4

You are going to read some information about schools. For questions 22-35 choose from the schools (A-H). Some of the schools may be chosen more than once. When more than one answer is required, these may be given in any order. There is an example at the beginning (0).

Which school(s) would you recommend for someone who:

- |  |    |                      |    |                      |
|--|----|----------------------|----|----------------------|
| is interested in French wines?                           | 0  | E                    |    |                      |
| is interested in computer programming?                   | 22 | <input type="text"/> |    |                      |
| wants to improve their writing skills?                   | 23 | <input type="text"/> |    |                      |
| enjoys learning about cookery?                           | 24 | <input type="text"/> |    |                      |
| wants to learn shorthand?                                | 25 | <input type="text"/> |    |                      |
| is interested in competing for something?                | 26 | <input type="text"/> | 27 | <input type="text"/> |
| wants to improve their computer skills?                  | 28 | <input type="text"/> | 29 | <input type="text"/> |
| wants to take the opportunity to go abroad?              | 30 | <input type="text"/> | 31 | <input type="text"/> |
| is interested in studying the language used by the deaf? | 32 | <input type="text"/> | 33 | <input type="text"/> |
| wants to paint outdoors?                                 | 34 | <input type="text"/> |    |                      |
| is interested in costume design?                         | 35 | <input type="text"/> |    |                      |



## Schools

### Pembrey Memorial Community Centre

A

Pembrey Memorial Community Centre continues its tradition of training in the domestic crafts, but this year, for the first time, Pembrey is also offering courses in arts and athletics. Monday, Tuesday and Thursday are reserved for cooking, embroidery and needlework classes. On Wednesdays we offer a break for those who wish to relax and exercise; golf or aerobics classes. There is also a new landscape painting class on offer. The tree-lined hills surrounding the campus are a beautiful area to paint.

### Glan-y-mor Comprehensive

B

Glan-y-mor Comprehensive school offers two exciting new courses this semester, both with guest instructors. A dressmaking course for those interested in designing clothes for the theatre will be offered on Wednesday nights. Teaching it is Madeline Albright, who has won numerous awards for her costumes in London theatres. She has done the dressmaking for Shakespeare's plays as well as contemporary musicals such as *Guys and Dolls*. Also offered for the first time is a course in computer skills. On Wednesdays the focus is on programming, while Thursdays are reserved for beginners wishing to learn word processing and other basic skills. Donald Matthews, a consultant with IBM in New York, teaches.

### Wales Centre

C

Because of a lack of sign language courses on offer in Wales, WC has decided to work with the Llanelli School for the Deaf to train teachers, parents and friends in this skill. Of course we're also proud to announce our tenth year in woodwork and pottery. For the occasion, a competition will be held at the end of the year for best design and construction in both wood and clay.

### Pentip Language School

D

The Pentip Language School teaches German, Spanish and Welsh at all levels. The classes are only offered on Mondays and Tuesdays this year. There's a special two-week programme in the Spanish department during the

spring semester. Twelve students will be able to travel to Madrid where they will attend classes, live with Spanish families and be placed in a Spanish company for three working days to observe and learn from the business community in Spain.

### The Community Education Centre

E

This year the Community Education Centre combines its French and wine-making courses to give an overall study of the French language and the geography of its rich wine country. There will be guest speakers from Bordeaux to explain the techniques used by some of their chateaux. Weekend trips to various regions of France can be arranged through the school.

### Lane Secretarial School

F

Our secretarial school has the highest job placement record of any similar school in Wales. We place 97 per cent of our students in various jobs throughout the country and abroad. Although classes are only offered on Tuesday evenings, every area is covered, including word processing, typing and shorthand. Beginner classes in French and Welsh have also been added to the programme.

### Laverne Beauty School

G

This year, Katy Decharise from L'Oreal will be teaching our make-up and beauty class. The class will be held on Friday evenings, but an additional session will take place on Saturdays for those interested in hairstyling and massage. There is also a new silk screen painting class, and a competition for the best T-shirt will be held at the end of each semester.

### Avenue Learning Centre

H

The Avenue Learning Centre is for those who need extra help with their reading, writing and communication skills. You can study maths for work or home, work on study skills or gain support while studying. Teachers are in the office at all hours to give private tutoring or answer any questions. For the first time, we are also offering a course in sign language for the hearing-impaired.

## PAPER 3 • USE OF ENGLISH

### Part 1

For questions 1 - 15, read the text below and decide which word **A, B, C** or **D** best fits each space. There's an example at the beginning (0). Mark your answers in the **answer boxes** provided.



### Relax

The concept of stress has been around for a long time, but recently people have been (0) ... about it, and suffering (1) ... it, more and more. Although it is often (2) ... as a psychological problem, this disorder (3) ... people's working lives as (4) ... as physical illness does. The first (5) ... is tiredness, followed by the (6) ... to sleep, problems with concentration and loss of (7) ... . None of these symptoms surprise Dr Koukopoulos, a psychiatrist working in Rome, who comments, "With (8) ... phones, faxes and home computers, people simply never stop. Life is now a thousand (9) ... more intense than it was 100 years ago. The human (10) ... can't adapt that quickly." One of the problems in dealing with stress is that sufferers don't discuss it for (11) ... of seeming a failure. This can lead to the use of anti-depressants and other drugs. This type

of (12) ... is not the best remedy. As an (13) ... way of reducing stress, companies all over Europe are introducing workshops, group therapy and relaxation (14) ... for their employees. The first step is to (15) ... the problem, the second is to talk about it and the last is to rebalance your life by learning to relax.

- |    |                 |              |              |              |
|----|-----------------|--------------|--------------|--------------|
| 0  | A discussing    | B talking    | C saying     | D telling    |
| 1  | A by            | B with       | C of         | D from       |
| 2  | A explained     | B believed   | C seen       | D given      |
| 3  | A interests     | B effects    | C involves   | D affects    |
| 4  | A much          | B far        | C long       | D well       |
| 5  | A symbol        | B sign       | C sight      | D signal     |
| 6  | A impossibility | B difficulty | C disability | D inability  |
| 7  | A hunger        | B taste      | C appetite   | D starvation |
| 8  | A mobile        | B moveable   | C moving     | D travelling |
| 9  | A times         | B ways       | C types      | D kinds      |
| 10 | A structure     | B body       | C frame      | D skeleton   |
| 11 | A anxiety       | B alarm      | C fear       | D worry      |
| 12 | A healing       | B management | C cure       | D treatment  |
| 13 | A alternative   | B opposite   | C other      | D alternate  |
| 14 | A meetings      | B terms      | C sessions   | D semesters  |
| 15 | A meet          | B recognise  | C notice     | D know       |

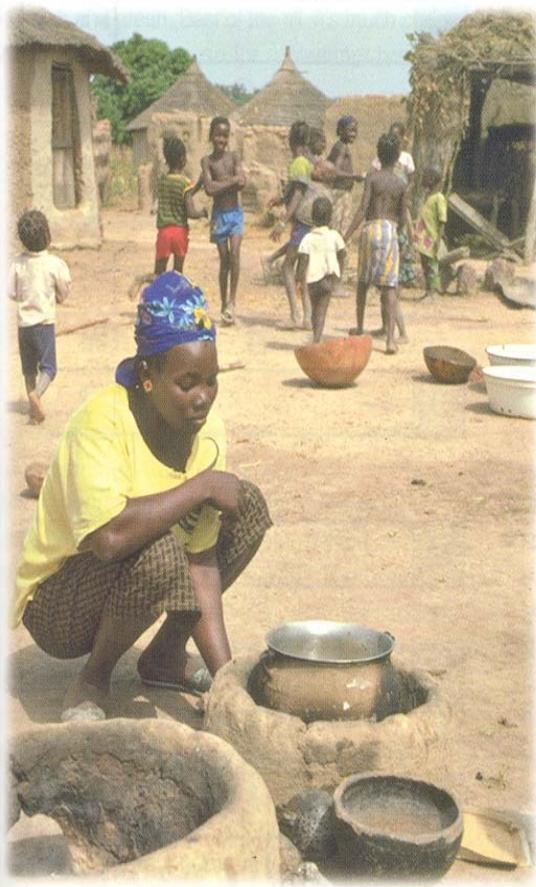
0	A	B	C	D
1	A	B	C	D
2	A	B	C	D
3	A	B	C	D
4	A	B	C	D
5	A	B	C	D
6	A	B	C	D
7	A	B	C	D
8	A	B	C	D
9	A	B	C	D
10	A	B	C	D
11	A	B	C	D
12	A	B	C	D
13	A	B	C	D
14	A	B	C	D
15	A	B	C	D

Part 2

For questions 16 - 30, read the text below and think of the word which best fits each space. Use only **one** word in each space. There is an example at the beginning (0). Write your answers in the **answer boxes** provided.

**The Population Problem**

As we near the end of the 20th Century, the world's population figure stands at (0) ..... to 6 billion and is expected to double during the next hundred years. However, only 10% of this growth is expected to occur (16) ..... rich, developed countries. This means that many poor, developing countries will find it impossible to provide everyone (17) ..... the basics needed (18) ..... survival - housing, food and fuel. Fields are already (19) ..... overgrazed by animals and overworked by farmers, and trees are being chopped down for fuel. All these things (20) ..... led to drought and famine in many parts of the world, particularly in Africa. Indian governmental reports state that 37% of its people are unable to buy enough food to sustain (21) ..... . As a (22) ....., officials fear the establishment of a new sub-human race of people (23) ..... are physically too weak and hungry to deal with their problems. To (24) ..... prevent this from happening, scientists want to (25) ..... the world's population growth in half over the (26) ..... ten years. Family planning and educational campaigns will be the main focus of this approach. Another possible solution (27) ..... combine family planning methods with some form of wealth redistribution. For many countries, (28) ....., the problem lies in the fact that they have so (29) ..... desperately poor citizens in the first (30) ..... . As the countries of the west have illustrated, the wealthier your population is, the slower it will grow.



0	close	0
16		16
17		17
18		18
19		19
20		20
21		21
22		22
23		23
24		24
25		25
26		26
27		27
28		28
29		29
30		30

## Part 3

For questions 31 - 40, complete the second sentence so that it has a similar meaning to the first sentence. Use the word given and other words to complete each sentence. You must use between **two and five** words. **Do not change the word given.** There is an example at the beginning (0). Write **only** the missing words in the **answer boxes** provided.

0 I'm sure it wasn't Jim who did it.  
**have**  
It ..... Jim who did it.

0	<i>can't have been</i>	0 0 1 2 □ □ □
---	------------------------	------------------

31 There wasn't anything he could say in light of the evidence.  
**nothing**  
There ..... in light of the evidence.

31		31 0 1 2 □ □ □
----	--	-------------------

32 Her parents insist that she phone them every day.  
**her**  
Her parents ..... every day.

32		32 0 1 2 □ □ □
----	--	-------------------

33 He never takes Sarah's problems seriously.  
**light**  
He ..... Sarah's problems

33		33 0 1 2 □ □ □
----	--	-------------------

34 It takes eight hours to fly to Saudi Arabia.  
**flight**  
It ..... to Saudi Arabia.

34		34 0 1 2 □ □ □
----	--	-------------------

35 This watch belonged to my grandfather.  
**was**  
This ..... watch.

35		35 0 1 2 □ □ □
----	--	-------------------

36 He didn't have time to type the report.  
**round**  
He didn't ..... the report.

36		36 0 1 2 □ □ □
----	--	-------------------

37 You may borrow only three books at one time.  
**be**  
Only three books ..... at one time.

37		37 0 1 2 □ □ □
----	--	-------------------

38 The children should watch their behaviour at the dinner table.  
**themselves**  
The children ..... at the dinner table.

38		38 0 1 2 □ □ □
----	--	-------------------

39 There are only a few people staying at this hotel.  
**not**  
There ..... staying at this hotel.

39		39 0 1 2 □ □ □
----	--	-------------------

40 Everyone was invited to the party except Julia.  
**person**  
The ..... invited to the party was Julia.

40		40 0 1 2 □ □ □
----	--	-------------------

## Part 4

For questions 41 - 55, read the text below and look carefully at each line. Some of the lines are correct and some have a word which should not be there.

If a line is correct, put a tick (✓) by the number in the **answer boxes** provided. If a line has a word which should **not** be there, write the word in the **answer boxes** provided. There are two examples at the beginning (0 and 00).

### Getting Away

0 Last summer it was so hot and humid in the London  
 00 that by August I felt I really couldn't bear it  
 41 any far longer. I had been planning to take my  
 42 holiday in September because of we were very busy at  
 43 work, but I was so tired out of the heat that I asked  
 44 for a couple of days off, threw some clothes into  
 45 a suitcase and took off the train to Cromer, which is on  
 46 the coast of Norfolk. Many people think about that Cromer  
 47 is too touristy to be attractive, but I found out it full  
 48 of beauty and charm. It's true there were a lots of  
 49 people there, but the beach is so big that by  
 50 walking for ten minutes you can escape all the  
 51 crowds. There are also some too lovely walks along the  
 52 tops of the cliffs, with magnificent views along  
 53 the coast, and the countryside inland it is very  
 54 lush and green. Best of the all, it's much cooler there  
 55 than in London - so for a pleasant break, try Cromer.

0	the	0
00	✓	00
41		41
42		42
43		43
44		44
45		45
46		46
47		47
48		48
49		49
50		50
51		51
52		52
53		53
54		54
55		55

## Part 5

For questions 56 - 65, read the text below. Use the word given in capitals at the end of each line to form a word that fits in the space in the same line. There is an example at the beginning (0). Write your word in the **answer boxes** provided.

### What are feelings?

The question of what love and hate actually are still remains (0) .... Is it simply a matter of (56) ....., or do emotions come from some unexplored (57) .... of our being which (58) .... have not yet discovered? The brain is the most (59) .... part of all living creatures, and that is where (60) .... are looking for the answers to these (61) .... questions. But up to now little (62) .... has been found to support the (63) .... theories that have been put forward. Perhaps this (64) .... into the nature of feelings will prove (65) ....., but such questions will continue to interest humans forever.

ANSWER  
 CHEMIST  
 DEEP  
 SCIENCE  
 MYSTERY  
 RESEARCH  
 PUZZLE  
 EVIDENT  
 VARY  
 EXPLORE  
 SUCCESSFUL

0	unanswered	0
56		56
57		57
58		58
59		59
60		60
61		61
62		62
63		63
64		64
65		65

## PAPER 4 • LISTENING

### Part 1

You will hear people talking in eight different situations. For questions 1 - 8, choose the best answer, **A**, **B** or **C**.

- 1** Listen to this man talking to someone at the check-in desk in Gatwick Airport. Why did he miss his flight to Glasgow?

**A** He lost his ticket.  
**B** Airport officials were on strike.  
**C** He was put on the wrong flight in Rome.

	1
--	---

- 2** Listen to this conversation between Lucy and her mother in a shop. What does Lucy want?

**A** to try on a long white dress  
**B** to have a fairytale wedding dress  
**C** to wear a fitted, cream dress at her wedding

	2
--	---

- 3** Listen to the conversation between Susan and her friend. Where is Susan going?

**A** to the dentist for some painkillers  
**B** to the pharmacy for some painkillers  
**C** to the dentist to have a tooth extracted

	3
--	---

- 4** Listen to these two friends talking about old family possessions. What did the woman do with her old books?

**A** threw them away  
**B** gave them away  
**C** sold them

	4
--	---

- 5** You hear a man talking on the telephone. Who is he phoning?

**A** an estate agent  
**B** a travel agent  
**C** a finance company

	5
--	---

- 6** You hear a man asking for advice from a wine merchant. Which wine does the man suggest?

**A** sweet white wine from France  
**B** sweet white wine from Santorini  
**C** dry white wine from Greece

	6
--	---

- 7** Listen to the shop assistant serving a woman. Which shoes do they have?

**A** size 40 in blue, size 39 in black  
**B** size 40 in blue, size 38 in black  
**C** size 39 in blue, size 38 in black

	7
--	---

- 8** Listen to this man phoning a workman about a job on his house. Why is he angry?

**A** The workman changed the price of the job.  
**B** The workman worked with a different builder.  
**C** The workman left before the job was finished.

	8
--	---

**Part 2**

You will hear a woman being interviewed for a job. For questions 9 - 18, fill in the application form.



**JOB APPLICATION FORM**

Post applied for:  9

College/University attended:  10

Qualification(s):  11

Age at Graduation:  12

First job:  13

Length of employment at previous job:  14

Reason for leaving:  15

Interest(s):  16

Special skills:  17

and

18

### Part 3

You will hear five people talking about beds. For questions 19 - 23, choose from the list **A - F** what each speaker is describing. Use the letters only once. There is one extra letter which you do not need to use.

**A** bunk bed

**B** single bed

**C** camp bed

**D** sofa bed

**E** water bed

**F** antique bed

Speaker 1

<input type="text"/>	19
----------------------	----

Speaker 2

<input type="text"/>	20
----------------------	----

Speaker 3

<input type="text"/>	21
----------------------	----

Speaker 4

<input type="text"/>	22
----------------------	----

Speaker 5

<input type="text"/>	23
----------------------	----



**Part 4**

You will hear a conversation between three friends who are organising a surprise party for a fourth friend.

Answer questions **24 - 30**, by writing

**G** (for Gabriel)

**E** (for Evelyn)

or **S** (for Steven)

in the boxes provided.

**24** Who had forgotten about Basil's birthday?

	<b>24</b>
--	-----------

**25** Who has had exams?

	<b>25</b>
--	-----------

**26** Who decides to help organise the music?

	<b>26</b>
--	-----------

**27** Who has invited Sam?

	<b>27</b>
--	-----------

**28** Who thinks they should buy more soft drinks?

	<b>28</b>
--	-----------

**29** Whose sister can't come?

	<b>29</b>
--	-----------

**30** Who has to meet someone at the train station?

	<b>30</b>
--	-----------





PRACTICE TEST 2

TEST OF WRITTEN ENGLISH (TWE)  
ESSAY QUESTION

Time – 30 minutes

DO NOT TURN THE PAGE UNTIL THE SUPERVISOR TELLS YOU TO DO SO.

The TWE essay question follows. You will have 30 minutes to plan, write, and correct your essay. Your essay will be graded on its overall quality.

1. When the Supervisor tells you to begin, turn the page and read the Essay Question carefully.
2. Think before you write. Making notes may help you to organize your essay. Below the Essay Question is a space for notes. Use only this area to outline your essay or make notes.
3. Write only on this topic. If you write an essay on a different topic, it will not be scored. Write clearly and precisely. Use examples to support your ideas. How well you write is much more important than how much you write, but to cover the topic adequately, you may want to write more than one paragraph.
4. Start writing your essay on the first line of Essay Page Side 3. Use Side 4 if you need more space. Extra paper will not be provided. Write neatly and legibly. Do not skip lines. Do not write in very large letters or leave large margins.
5. Check your work. Allow a few minutes before time is called to read over your essay and make small changes.
6. After 30 minutes, the Supervisor will tell you to stop. You must stop writing and put your pencil down. If you continue to write, it will be considered cheating.

STOP! WAIT FOR THE SUPERVISOR'S INSTRUCTIONS.



**SECTION 1**  
**LISTENING COMPREHENSION**

In this section of the test, you will have an opportunity to demonstrate your ability to understand spoken English. There are three parts to this section, with special directions for each part. Do not read ahead or turn the pages while the directions are being read. Do not take notes or write in your workbook at any time.

**Part A**

Directions: For each question in Part A, you will hear a short sentence. Each sentence will be spoken just once. The sentences you hear will not be written out for you.

After you hear a sentence, read the four choices in your workbook, marked (A), (B), (C), and (D), and decide which one is closest in meaning to the sentence you heard. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen. Fill in the space completely so that the letter inside the oval cannot be seen.

**Listen to an example.**

On the recording, you hear: *The new observatory is nothing less than outstanding.*

- In your workbook, you read:
- (A) The new observatory is excellent.
  - (B) There is nothing outstanding to observe.
  - (C) You must stand outside to make the observance.
  - (D) His new observations are more outstanding than usual.

**Sample Answer**



The man said, "The new observatory is nothing less than outstanding." Sentence (A), "The new observatory is excellent," is closest in meaning to what the man said. Therefore, the correct choice is (A).

NOW WE WILL BEGIN PART A WITH QUESTION 1.



1. (A) Peter asked the waiter to bring in his friend.  
(B) Peter asked the waiter for the bill.  
(C) Bill and Peter asked the waiter for a table.  
(D) Peter ordered a meal from the waiter.
2. (A) David had a good solution for his friend.  
(B) David solved his friend's problems.  
(C) A good friend of David's showed him the problems.  
(D) A good friend found an answer to David's problems.
3. (A) We couldn't finish on time for Joyce.  
(B) Joyce enabled us to meet her in time.  
(C) We were capable of meeting Joyce's train.  
(D) Joyce wasn't able to save our place in line.
4. (A) If Diana wants to wash the car, she will.  
(B) Let Diana wash the car, since she is willing.  
(C) Since you asked Diana to wash the car, she will.  
(D) Why don't you ask Diana to wash the car?
5. (A) Lisa's wallet must be at my place.  
(B) Lisa doesn't know where her billfold is.  
(C) Lisa's billfold was lost on the bus.  
(D) Lisa was put in the wrong place.
6. (A) Since you were in New York, you must have visited the Statue of Liberty.  
(B) If you go to New York, you should visit the Statue of Liberty.  
(C) I think you should have visited the Statue of Liberty while you were in New York.  
(D) Did you visit the Statue of Liberty while you were in New York?
7. (A) We don't have to buy a new Ford.  
(B) We don't have enough money to buy a new car.  
(C) We shouldn't offer to buy a new car.  
(D) We can't buy four new cars!
8. (A) If you can't see the point, tell me.  
(B) If you don't understand, you should ask.  
(C) If you're unable to come, please let me know.  
(D) You should shake the ointment, if you ask me.
9. (A) Has she really earned that many points?  
(B) I didn't realize she'd put on so much weight.  
(C) It is true that she can sell grain?  
(D) She really needs to gain weight.
10. (A) People are cheerful at our house when they return from the stadium.  
(B) The stadium has clouds over it on most days.  
(C) The noise from the stadium is too loud.  
(D) The stadium is near enough to our house that you can hear the people shouting.
11. (A) The flight attendant and the passengers got on the airplane.  
(B) The flight attendant greeted the passengers as they boarded the plane.  
(C) The flight attendant thanked the passengers for getting on board.  
(D) The bored passengers were met by the flight attendant of the plane.



12. (A) There's nothing certain about Jane's success.  
(B) Jane's determination has nothing to do with her success.  
(C) Jane certainly is determined to be successful.  
(D) Jane's success has never been determined for certain.
13. (A) Mary only passed this test because she scored higher than the others.  
(B) Mary's score was barely enough to pass.  
(C) Only on this test was Mary's score high enough.  
(D) Mary was the only one to pass this test.
14. (A) Susan, like Scott, finds it difficult to study mathematics.  
(B) Mathematics is harder for Susan than for Scott.  
(C) Susan studied less than Scott for the math class.  
(D) The math class was harder for Susan than for Scott.
15. (A) The newspaper that came is mine!  
(B) Shouldn't the newspaper be here by now?  
(C) Please come and check the newspaper.  
(D) Please find out whether or not the newspaper is here.
16. (A) They didn't want to stay home or go to the movies.  
(B) They prefer staying home to going to the movies.  
(C) They don't like to go to the movies.  
(D) They prefer watching movies at home to going out.
17. (A) Betty was very late.  
(B) It was a two-hour wait for Betty.  
(C) Betty was over at Elliott's for two hours.  
(D) Elliott hates to be at Betty's for more than two hours.
18. (A) William sold three of them.  
(B) William isn't very well yet.  
(C) William's language is advanced.  
(D) William was not quiet when he was younger.
19. (A) We were used to playing tennis.  
(B) Now we like to play tennis.  
(C) We've been playing tennis longer than we've been watching it.  
(D) We stopped playing tennis when we got injured.
20. (A) They forgot to charge for breakfast.  
(B) Guests at the hotel must pay additionally for breakfast.  
(C) Breakfast is not served at the hotel.  
(D) Do not pay for breakfast at that hotel.

GO ON TO THE NEXT PAGE 



### Part B

Directions: In Part B you will hear short conversations between two people. After each conversation, a third person will ask a question about what was said. You will hear each conversation and question about it only one time. After you hear a conversation and the question about it, read the four possible answers in your workbook and decide which one is the best answer to the question you heard. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen.

#### Listen to an example.

On the recording, you hear: (Man) *I think I'll have the curtains changed.*  
(First Woman) *They are a bit worn.*  
(Second Woman) *What does the woman mean?*

In your workbook, you read: (A) She thinks every bit of change is important.  
(B) She wants to wear them.  
(C) She thinks they've been worn enough.  
(D) She thinks they are in bad condition.

#### Sample Answer

(A) (B) (C) ●

You learn from the conversation that the woman thinks the curtains are worn. The best answer to the question "What does the woman mean?" is (D), "She thinks they are in bad condition." Therefore, the correct choice is (D).

NOW WE WILL BEGIN PART B WITH THE FIRST CONVERSATION.





21. (A) The students asked people to make them.  
(B) The students asked people to buy them.  
(C) The students asked people to post them.  
(D) The students asked people to contribute them.
22. (A) Making salad.  
(B) Tossing balls.  
(C) Playing a board game.  
(D) Designing squares.
23. (A) He couldn't get a ticket.  
(B) He bought a ticket weeks ago.  
(C) He reserved a ticket for tomorrow.  
(D) The concert has been canceled.
24. (A) Mr. Brown isn't good at repairing TVs.  
(B) The TV doesn't need fixing.  
(C) The TV cannot be fixed.  
(D) He can repair the TV himself.
25. (A) She's always hungry.  
(B) She's full of apples.  
(C) She's on a diet.  
(D) She's pleased he packed them.
26. (A) Nancy will want to write her own notes.  
(B) She can't imagine Nancy writing notes.  
(C) She doesn't want to lend her notes to Nancy.  
(D) She thinks Nancy won't be able to read the notes.
27. (A) Tom and Lucy getting standing-room-only tickets.  
(B) Tom and Lucy stopping traffic.  
(C) Tom and Lucy being stuck in traffic.  
(D) Tom and Lucy standing outside in the snow.
28. (A) Jennifer has to look for a spring.  
(B) Jennifer must learn those facts in order to graduate.  
(C) Jennifer has faced many problems this spring.  
(D) Jennifer must accept not being able to graduate.
29. (A) She set the pail on the flowers.  
(B) She bought light blue writing paper.  
(C) She had to water the flowers.  
(D) She got the blue flowers behind the station.
30. (A) He is indifferent about the chemistry class.  
(B) He doesn't mind the chemistry class.  
(C) He isn't careful in chemistry class.  
(D) He doesn't like his chemistry lessons.
31. (A) Her mother likes the potato salad.  
(B) Her mother makes better potato salad.  
(C) Her mother says there's never enough potato salad.  
(D) Her mother has never tasted the potato salad.
32. (A) Sam doesn't like parties.  
(B) Sam wishes he weren't at the party.  
(C) Sam's personality makes parties more enjoyable.  
(D) Sam lives above the place where the party is.
33. (A) During Christmas.  
(B) The last week before Christmas.  
(C) About two weeks ago.  
(D) Last week.
34. (A) The man is blind.  
(B) It is snowing very hard.  
(C) She doesn't know if it's snowing.  
(D) She can't get across the street.
35. (A) He called for the string.  
(B) He cut his finger.  
(C) He fingered the tie.  
(D) He remembered the box.

GO ON TO THE NEXT PAGE 



### Part C

Directions: In this part of the test, you will hear longer conversations and talks. After each conversation or talk, you will be asked some questions. You will hear the talks and conversations and the questions about them only one time. They will not be written out for you.

After you hear a question, read the four possible answers in your workbook and decide which one is the best answer to the question you heard. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen. Answer all questions on the basis of what is stated or implied by the speakers in the talk or conversation.

**Here is an example.**

On the recording, you hear: *Listen to a conversation between a professor and a student.*

(First Woman) *Would you please type this paper? I can't read your handwriting.*

(Man) *I'm sorry, Professor Mills. I don't have a typewriter, and besides, I can't type.*

(First Woman) *Well, rewrite it then, but be sure to make it clear.*

**Now listen to a sample question.**

(Second Woman) *What does the professor want the man to do?*

- In your workbook, you read:
- (A) Buy a typewriter.
  - (B) Read his handwriting.
  - (C) Clear the papers away.
  - (D) Rewrite the paper.

**Sample Answer**

(A) (B) (C) (D)

The best answer to the question "What does the professor want the man to do?" is (D), "Rewrite the paper." Therefore, the correct choice is (D).

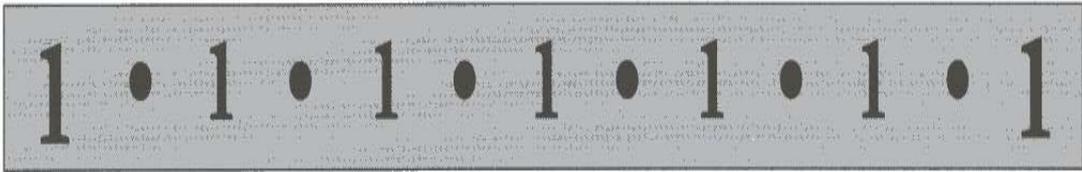
Remember, you are not allowed to take notes or write in your workbook.

NOW WE WILL BEGIN PART C.

GO ON TO THE NEXT PAGE 

36. (A) The library.  
(B) A university classroom.  
(C) The reserve desk.  
(D) The microfiche room.
37. (A) They are no longer published.  
(B) They're on microfiche.  
(C) They cannot be found in the library.  
(D) They must be purchased at the bookstore.
38. (A) They cannot be taken out of the library.  
(B) They are out of print.  
(C) They have an asterisk.  
(D) They must be read before the following class.
39. (A) The librarian can teach her.  
(B) She can read the microfiche.  
(C) The instructions are on the folder.  
(D) There are instruction pamphlets in the machine room.
40. (A) 100 monkeys on an island.  
(B) Islands off the coast of Japan.  
(C) A study showing the interconnectedness of life.  
(D) The Japanese sweet potato monkey.
41. (A) Copy the behavior of other monkeys.  
(B) Eat sweet potatoes in a particular way.  
(C) Contact monkeys 200 miles away.  
(D) Teach monkeys on other islands.
42. (A) Japan.  
(B) 200 miles from Japan.  
(C) An island off Japan.  
(D) A sweet potato farm.
43. (A) It was not typical of monkeys.  
(B) The other monkeys refused to learn it.  
(C) It was difficult to learn.  
(D) Monkeys learned it at different rates.
44. (A) At Mary's.  
(B) At Jeff's.  
(C) At Mary's parents'.  
(D) At Jennifer and Lisa's.
45. (A) Lisa.  
(B) Jennifer.  
(C) Mary's parents.  
(D) Jeff.
46. (A) They've never met each other.  
(B) They knew Mary's parents.  
(C) They live together.  
(D) They need Jeff's help.

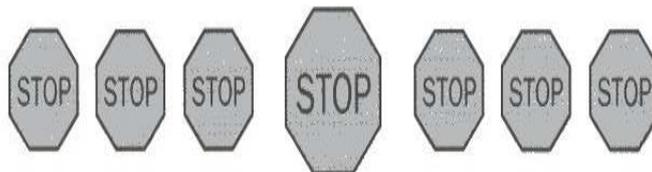
GO ON TO THE NEXT PAGE 



47. (A) They are roommates.  
(B) They are married.  
(C) They are classmates.  
(D) They are friends.
48. (A) Improving your ability to perform sports skills.  
(B) Carrying through a sport-specific motion.  
(C) Ingraining patterns in your subconscious movement memory.  
(D) Improving arm muscles for better performance.
49. (A) Repeating the movement.  
(B) Memorizing the pattern.  
(C) Performing with less effort.  
(D) Applying greater speed.
50. (A) It is less important than strengthening the larger muscles.  
(B) It is time consuming.  
(C) It reduces your risk of injury.  
(D) It is not part of most athletes' fitness program.

**THIS IS THE END OF THE LISTENING COMPREHENSION SECTION OF PRACTICE TEST 2.**

THE NEXT PART OF THE TEST IS SECTION 2. TURN TO THE  
DIRECTIONS FOR SECTION 2 IN YOUR WORKBOOK,  
READ THEM, AND BEGIN WORK.  
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.



**SECTION 2**  
**STRUCTURE AND WRITTEN EXPRESSION**

Time – 25 minutes

This section is designed to measure your ability to recognize language that is appropriate for standard written English. There are two types of questions in this section, with special directions for each type.

Directions: Questions 1–15 are incomplete sentences. Beneath each sentence you will see four words or phrases marked (A), (B), (C), and (D). Choose the one word or phrase that best completes the sentence. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen. Fill in the space so that the letter inside the oval cannot be seen.

**Example** John Le Carré ----- for his novels concerning espionage.

- (A) famous
- (B) has fame
- (C) is famous
- (D) famed for

**Sample Answer**

(A) (B) ● (D)

The sentence should read, "John Le Carré is famous for his novels concerning espionage." Therefore, you should choose answer (C).

After you read the directions, begin work on the questions.

1. ----- infinitely large number of undiscovered galaxies.
  - (A) An
  - (B) There are an
  - (C) From an
  - (D) Since there are
2. The Great Wall of China is perhaps the most awe inspiring ----- the great structures of the world.
  - (A) that
  - (B) with
  - (C) among
  - (D) alone
3. No matter -----, Mozart was an accomplished composer while still a child.
  - (A) how remarkable it seems
  - (B) how it seems remarkable
  - (C) how seems it remarkable
  - (D) how it remarkable seems
4. Novelist Jane Austen ----- the first twenty-six years of her life in the village of Steventon, Hampshire.
  - (A) living
  - (B) was lived
  - (C) lived
  - (D) who lived
5. By far ----- of Saudi Arabia is oil.
  - (A) it is the most important export
  - (B) the most important export is
  - (C) that is the most important export
  - (D) the most important export
6. Yellowstone National Park's attractions include the famous Old Faithful geyser, vast forests, plentiful wildlife, and -----.
  - (A) campgrounds are well maintained
  - (B) well-maintained campgrounds
  - (C) campgrounds are maintained well
  - (D) maintains campgrounds well
7. ----- the Dartmoor ewe and the Dumpy hen, the Tamworth pig is a rare breed of farm animal.
  - (A) How
  - (B) Like
  - (C) The
  - (D) Even

**GO ON TO THE NEXT PAGE**

# 2 • 2 • 2 • 2 • 2 • 2 • 2 • 2

8. ----- often serve as places of public entertainment and festivals, they can also be places where people can find peace and solitude.  
 (A) Even though city parks  
 (B) City parks  
 (C) City parks that  
 (D) There are city parks which
9. ----- in the planet Uranus dates from its discovery in 1781 is not surprising.  
 (A) That scientific interest  
 (B) It was scientific interest  
 (C) Though scientific interest  
 (D) Scientific interest
10. The Himalaya Mountains are the -----.  
 (A) height of world extensive ranges  
 (B) ranges of the most extensive world  
 (C) world's most extensive ranges  
 (D) extensive ranges of the world
11. -----, Luxor did not reach preeminence until about 2,000 B.C.  
 (A) Many centuries earlier it was founded  
 (B) The city founded centuries earlier  
 (C) Although founded many centuries earlier  
 (D) Founding the city centuries earlier
12. Early sailing ships, ----- sometimes in uncharted seas, faced many hazards in reaching their destination.  
 (A) navigating  
 (B) were navigated  
 (C) navigate  
 (D) and navigates
13. Not until about a century after Julius Caesar landed in Britain ----- actually conquer the island.  
 (A) the Romans did  
 (B) did the Romans  
 (C) the Romans  
 (D) Romans that
14. ----- is as simple as it looks.  
 (A) Adobe houses building  
 (B) Adobe houses which are built  
 (C) Building adobe houses  
 (D) Adobe houses built
15. Square-rigged ships, ----- high speeds only when traveling with the trade winds, are not used commercially nowadays.  
 (A) can attain  
 (B) when attaining  
 (C) they can attain  
 (D) which can attain

Directions: In questions 16–40 each sentence has four underlined words or phrases. The four underlined parts of the sentence are marked (A), (B), (C), and (D). Identify the one underlined word or phrase that must be changed in order for the sentence to be correct. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen.

**Example** The balloonists who remained aloft in the air radioed the control center.  
 A B C D

**Sample Answer**

(A) (B) (C) (D)

The sentence should read, "The balloonists who remained aloft radioed the control center." Therefore, you should choose answer (C).

After you read the directions, begin work on the questions.

16. Drying food by means of solar energy is a ancient process applied wherever food and climatic conditions make it possible.  
 A B C D
17. An itch resulting when a nerve that can carry pain is only slightly stimulated.  
 A B C D
18. A healthy person snores most because the membrane in the nose becomes dry.  
 A B C D

GO ON TO THE NEXT PAGE 

# 2 • 2 • 2 • 2 • 2 • 2 • 2 • 2

19. For they, the most rewarding shipwreck found was the HMS Associated.  
 A B C D
20. America's first satellites exploded before it had risen three and a half feet off the ground.  
 A B C D
21. The Concorde can fly across the Atlantic without refueling and carrying 11 tons of freight.  
 A B C D
22. Most babies will grow up to be as cleverer as their parents.  
 A B C D
23. With animals both humans, chewing helps relieve tension.  
 A B C D
24. Diane Arbus's unusual and controversial work includes photograph of sixties celebrities.  
 A B C D
25. It should not be assume that the lower the price, the happier the buyer.  
 A B C D
26. To help policymakers and another, the U.S. government spends as much as \$1.4 billion a year in collecting statistics.  
 A B C D
27. The Victorian constructions of Haight-Ashbury are among the fewer architectural survivors of the San Francisco earthquake in 1906.  
 A B C D
28. One out of every eight balloons in the world are launched at Albuquerque, New Mexico.  
 A B C D
29. At space camp, youngsters go through concentrated astronaut training but shuttle simulations.  
 A B C D
30. Sprinkler system have proven to be the most effective means of fighting hotel fires.  
 A B C D
31. For more than 450 years, Mexico City has been the economic, culture, and political center of the Mexican people.  
 A B C D

GO ON TO THE NEXT PAGE 

# 2 • 2 • 2 • 2 • 2 • 2 • 2 • 2

32. The rings of Saturn are so distant to be seen from Earth without a telescope.  
A B C D
33. In 1927, Charles Lindbergh completed a solo crossing of the Atlantic Ocean in its  
A B C D  
aircraft, the *Spirit of St. Louis*.
34. There is estimated that the Orion nebula contains enough matter to form 10,000  
A B C  
stars.  
D
35. Lack of animal protein in the human diet is a serious cause of the malnutrition.  
A B C D
36. Measle has not yet been eradicated because of controversy concerning  
A B C D  
immunization.
37. The firing of bricks and tiles for use in the build industry requires large amounts  
A B C D  
of fuel.
38. The cinecamera has fastly become obsolete with the invention of video recording  
A B C D  
technology.
39. Plants known so adaptogens improve human vitality, eliminate stress, and sustain  
A B C  
life in the most arduous conditions.  
D
40. A new battery for domestic appliances in the house is as thin and flexible as paper.  
A B C D

**THIS IS THE END OF SECTION 2 OF PRACTICE TEST 2.**

IF YOU FINISH BEFORE TIME IS UP, CHECK YOUR WORK  
ON SECTION 2 ONLY.  
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.



**SECTION 3**  
**VOCABULARY AND READING COMPREHENSION**

Time – 45 minutes

This section is designed to measure your comprehension of standard written English. There are two types of questions in this section, with special directions for each type.

Directions: In questions 1–30 each sentence has an underlined word or phrase. Below each sentence are four other words or phrases, marked (A), (B), (C), and (D). You are to choose the one word or phrase that best keeps the meaning of the original sentence if it is substituted for the underlined word or phrase. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter you have chosen. Fill in the space so that the letter inside the oval cannot be seen.

**Example** Many women have made a mark on American art since the early 1800s.

- (A) target for
- (B) dedication to
- (C) contribution to
- (D) degree of

Sample Answer

(A) (B) ● (D)

The best answer is (C) because “Many women have made a contribution to American art since the early 1800s” is closest in meaning to the original sentence. Therefore, you should choose answer (C).

After you read the directions, begin work on the questions.

1. Cast iron covers for street holes were originally decorated to provide sure footing for horses.
  - (A) real
  - (B) firm
  - (C) kind
  - (D) fine
2. Coral needs strong currents that keep the water temperature around it ideal for its growth.
  - (A) waves
  - (B) tides
  - (C) pressures
  - (D) streams
3. In Venice even the buildings of the common quarters are often beautified with small carvings.
  - (A) created
  - (B) described
  - (C) adorned
  - (D) engineered
4. Snake venom can be vacuum dried and used for the production of antivenin serum.
  - (A) potion
  - (B) saliva
  - (C) poison
  - (D) flesh
5. The root system of the water hyacinth acts as a filter by removing suspended specks from the water.
  - (A) portions
  - (B) beams
  - (C) debris
  - (D) particles
6. The whale’s inexplicable predilection for beaching itself is the second greatest threat to its survival.
  - (A) unexplainable
  - (B) unusual
  - (C) unforgettable
  - (D) essential
7. The use of flags as symbols of national identity began to develop a thousand years ago.
  - (A) signs
  - (B) plans
  - (C) codes
  - (D) goals
8. The problem was caused by a malfunction in the mechanism used to release the parachute.
  - (A) disengage
  - (B) relent
  - (C) liberate
  - (D) rescue

**GO ON TO THE NEXT PAGE** 

9. The microtelephone is a miniaturized telephone transmitter and receiver that can be inserted in the ear.  
(A) attached  
(B) impaled  
(C) heard  
(D) placed
10. Because animals such as hedgehogs, shrews and moles live on insects, they are very beneficial to humans.  
(A) distressing  
(B) helpful  
(C) forgiving  
(D) careful
11. Charles Wheatstone invented the concept of three-dimensional drawing in 1838.  
(A) model  
(B) condition  
(C) notion  
(D) thought
12. The task of note taking in Braille is fatiguing and time consuming.  
(A) reading  
(B) chore  
(C) stroke  
(D) practicing
13. Wheezing is thought to occur when there is a narrowing of airways within the lungs.  
(A) enveloping  
(B) respiration  
(C) agitation  
(D) constricting
14. When a magnet is immersed in liquid oxygen, its pulling power is intensified.  
(A) exposed  
(B) submerged  
(C) dropped  
(D) dissolved
15. George Gershwin gathered motifs for his opera *Porgy and Bess* while living in Charleston.  
(A) themes  
(B) support  
(C) talent  
(D) money
16. Bivalve mollusks are widely distributed in both marine and fresh water.  
(A) invested  
(B) detailed  
(C) dispersed  
(D) located
17. The arms race, conducted in the quest for security, has itself become a greater source of insecurity.  
(A) need  
(B) program  
(C) perception  
(D) search
18. Chromosomes usually adhere to each other in most cell stages.  
(A) adjust  
(B) move  
(C) cling  
(D) come
19. In today's technical world, language needs to be compatible with both humans and machines.  
(A) in accord with  
(B) translated to  
(C) compared with  
(D) disclosed to
20. Cellulose is the most abundant organic compound in nature.  
(A) varied  
(B) plentiful  
(C) discrete  
(D) disruptive
21. Many survivors of airplane crashes subsequently die because rescue teams have difficulty in locating the plane.  
(A) unfortunately  
(B) necessarily  
(C) usually  
(D) later
22. Milk production may have been the first purpose for which sheep were domesticated.  
(A) herded  
(B) housed  
(C) fenced  
(D) tamed
23. The polar bear's adaptation to its environment may have been achieved by combining a state of hibernation with one of physical activity.  
(A) system  
(B) condition  
(C) portion  
(D) position

24. The main routes used by the pony express were equipped with stops for food and lodging.  
 (A) stables  
 (B) shelter  
 (C) drinks  
 (D) sleep
25. Architects must take noise levels into consideration when designing an auditorium.  
 (A) into account  
 (B) into view  
 (C) seriously  
 (D) positively
26. Many dangerous areas had to be reconnoitered by pioneers crossing the Western plains.  
 (A) set aside  
 (B) entered into  
 (C) scouted out  
 (D) sorted out
27. Light hovercrafts can traverse very effectively virtually any surface.  
 (A) cross  
 (B) dig  
 (C) detour  
 (D) elevate
28. Legend has it that people known as the Mongulala were the forebears of the Incas.  
 (A) dependents  
 (B) enemies  
 (C) destroyers  
 (D) predecessors
29. The Florida Power and Light Company serves more than two million customers.  
 (A) visitors  
 (B) clients  
 (C) tourists  
 (D) settlers
30. Feelings of ambivalence are common during courtship.  
 (A) cowardliness  
 (B) recognition  
 (C) adoration  
 (D) uncertainty

---

**Directions:** In the rest of this section you will read several passages. Each one is followed by several questions about it. For questions 31–60, you are to choose the best answer, (A), (B), (C), or (D), to each question. Then, on your answer sheet, find the number of the question and fill in the space that corresponds to the letter of the answer you have chosen.

Answer all questions following a passage on the basis of what is stated or implied in that passage.

**Read the following passage.**

The horse has played a little-known but very important role in the field of medicine. Horses were injected with toxins of diseases until their blood built up immunities. Then a serum was made from their blood. Serums to fight both diphtheria and tetanus were developed in this way.

- Example** According to the passage, horses were given  
 (A) poisons from illnesses  
 (B) immunities to diseases  
 (C) diphtheria and tetanus serums  
 (D) medicines to fight toxins

**Sample Answer**

(B)  (C)  (D)

The passage states that “horses were injected with toxins of diseases.” Therefore, you should choose answer (A).

After you read the directions, begin work on the questions.

GO ON TO THE NEXT PAGE 

## Questions 31–35

- 1 The debt of lawn tennis to its French origins is illustrated in the unusual scoring  
2 system. This system probably stems from the habit of betting on individual points  
3 by the players or supporters. A game was worth one *denier*, so the points were  
4 worth the most convenient divisions of a *denier*. These were 15, 30, and 45 *sous*.  
5 In time, the latter became 40.  
6 Deuce, when both players have reached 40 in a game, is a corruption of the  
7 French *à deux*, meaning “both.” This may refer to both players having the same  
8 score – or to the fact that a player will need to take both the next two points to win.  
9 The term (as *dewce*) was first known in England in 1598.  
10 The word “love,” which means nil, may well come from the French word *l’oeuf*,  
11 meaning egg. The explanation for the use of the word *l’oeuf* is said to be the  
12 similarity of the shape of an egg to a zero. Modern player slang for a 6–0 6–0  
13 result is “egg and egg.”
31. What is the main topic of this passage?  
(A) betting and its effect on the game of tennis  
(B) the corruption of the French terminology for tennis  
(C) differences in terminology for the game of tennis in England and France  
(D) the influence of the French language on the scoring system of tennis
32. It can be inferred that the word “denier” in line 3 was  
(A) a monetary unit  
(B) a point system  
(C) a division  
(D) a score
33. The term “deuce”  
(A) comes from the English word “dewce”  
(B) means that the player scored two points  
(C) is used when both players have reached a score of 40  
(D) is a corruption of the French word meaning “two”
34. In line 10, the word “love” is used to refer to  
(A) tennis fans  
(B) two points  
(C) a score of zero  
(D) a winning point
35. An example of the corruption of French in the game of tennis is the word  
(A) love  
(B) nil  
(C) both  
(D) egg

## Questions 36–41

1 One of the major hazards for deep-sea divers is the “bends.” This condition is  
 2 caused by gas bubbles forming in the bloodstream if the diver ascends too rapidly.  
 3 The reason for this condition has to do with the saturation and desaturation of body  
 4 tissues with various gases. At increasingly greater depths, the diver breathes air at  
 5 higher pressures. This results in an increased quantity of air being dissolved in the  
 6 bloodstream. Different body tissues are saturated with different gases from the air  
 7 at different rates. When the diver ascends, oxygen is used by the body tissues,  
 8 carbon dioxide is released quickly, and nitrogen remains. The nitrogen needs to be  
 9 released gradually from the bloodstream and body tissues. If nitrogen is subjected  
 10 to a too rapid pressure reduction, it forms gas bubbles in the blood vessels. The  
 11 bubbles become trapped in the capillaries. This prevents blood and oxygen from  
 12 supplying necessary nutrients to body tissues, which consequently begin to die.  
 13 Saturation and desaturation are affected by various factors such as the depth,  
 14 length of time, and amount of exertion under water. There are other factors that a  
 15 diver must take into account when determining a safe ascent rate. These include  
 16 the diver’s sex and body build, the number of dives undertaken within the previous  
 17 12 hours, the time spent at the dive location before the dive, and the composition  
 18 of the respiration gas.

36. The passage is mainly about  
 (A) how to calculate a safe depth when diving  
 (B) how to determine saturation and desaturation rates  
 (C) instructions for diving safely  
 (D) the factors causing the bends in divers
37. It can be inferred from the passage that  
 (A) a woman is more likely to get the bends  
 (B) men and women may ascend at different rates  
 (C) men and women of certain athletic builds shouldn’t dive  
 (D) men are better divers than women
38. According to the passage, gas bubbles  
 (A) trap the capillaries  
 (B) are gradually released from the tissues and bloodstream  
 (C) block the supply of nutrients to body tissues  
 (D) are formed from compressed air
39. Which of the following does NOT affect the desaturation of body tissues?  
 (A) the location of the dive  
 (B) the number of previous dives  
 (C) the composition of the gas being used  
 (D) the amount of activity under water
40. According to the passage, the bends  
 (A) is the major diving hazard  
 (B) reduces pressure in the bloodstream  
 (C) is a condition caused by diving too quickly  
 (D) is a direct result of dying body tissues
41. According to the passage, which of the following is NOT true?  
 (A) Air at higher pressure is taken in at greater depths.  
 (B) More air is dissolved into the bloodstream at increasing depths.  
 (C) Carbon dioxide stays in the body when the diver ascends.  
 (D) Body tissues are saturated at different rates.


 GO ON TO THE NEXT PAGE

## Questions 42–47

- 1 Water scarcity is fast becoming one of the major limiting factors in world crop  
 2 production. In many areas, poor agricultural practices have led to increasing  
 3 desertification and the loss of formerly arable lands. Consequently, those plant  
 4 species that are well adapted to survival in dry climates are being looked at for an  
 5 answer in developing more efficient crops to grow on marginally arable lands.
- 6 Plants use several mechanisms to ensure their survival in desert environments.  
 7 Some involve purely mechanical and physical adaptations, such as the shape of  
 8 the plant's surface, smaller leaf size, and extensive root systems. Some of the  
 9 adaptations are related to chemical mechanisms. Many plants, such as cacti, have  
 10 internal gums and mucilages which give them water-retaining properties. Another  
 11 chemical mechanism is that of the epicuticular wax layer. This wax layer acts as  
 12 an impervious cover to protect the plant. It prevents excessive loss of internal  
 13 moisture. It also protects the plant from external aggression, which can come from  
 14 inorganic agents such as gases, or organic agents which include bacteria and plant  
 15 pests.
- 16 Researchers have proposed that synthetic waxes with similar protective abilities  
 17 could be prepared based on knowledge of desert plants. If successfully developed,  
 18 such a compound could be used to greatly increase a plant's ability to maintain  
 19 health in such adverse situations as inadequate water supply, limited fertilizer  
 20 availability, attack by pests, and poor storage after harvesting.
42. This passage deals mainly with  
 (A) desertification  
 (B) decreasing water supplies  
 (C) factors limiting crop production  
 (D) developing efficient plants
43. Which of the following is a mechanical or physical mechanism desert plants use?  
 (A) the plant's shape  
 (B) the small root system  
 (C) the vast leaf size  
 (D) the high water consumption
44. Which is one of the ways the epicuticular wax protects the plant?  
 (A) It helps the plant to avoid excessive moisture intake.  
 (B) It helps the plant to attack aggressors.  
 (C) It releases gases against plant pests.  
 (D) It guards against bacteria.
45. It can be inferred that synthetic simulated waxes have  
 (A) not been developed yet  
 (B) not succeeded  
 (C) been determined to be impervious to organic and inorganic agents  
 (D) the quality of causing bacteria
46. What is an example of an inorganic agent that may attack plants?  
 (A) bacteria  
 (B) insects  
 (C) gas  
 (D) pests
47. What is NOT an example of an adverse situation for crops?  
 (A) inadequate water  
 (B) insufficient fertilizer  
 (C) pest aggression  
 (D) proper storage

## Questions 48–53

1 Every year about two million people visit Mount Rushmore, where the faces of  
 2 four U.S. presidents were carved in granite by sculptor Gutzon Borglum and his  
 3 son, the late Lincoln Borglum. The creation of the Mount Rushmore monument  
 4 took 14 years – from 1927 to 1941 – and nearly a million dollars. These were  
 5 times when money was difficult to come by and many people were jobless. To  
 6 move the more than 400,000 tons of rock, Borglum hired laid-off workers from the  
 7 closed-down mines in the Black Hills area. He taught these men to dynamite, drill,  
 8 carve, and finish the granite as they were hanging in midair in his specially devised  
 9 chairs, which had many safety features. Borglum was proud of the fact that no  
 10 workers were killed or severely injured during the years of blasting and carving.

11 During the carving, many changes in the original design had to be made to keep  
 12 the carved heads free of large fissures that were uncovered. However, not all the  
 13 cracks could be avoided, so Borglum concocted a mixture of granite dust, white  
 14 lead, and linseed oil to fill them.

15 Every winter, water from melting snows gets into the fissures and expands as it  
 16 freezes, making the fissures bigger. Consequently, every autumn maintenance work  
 17 is done to refill the cracks. The repairers swing out in space over a 500-foot drop  
 18 and fix the monument with the same mixture that Borglum used to preserve this  
 19 national monument for future generations.

48. In line 12, the word “fissures” refers to  
 (A) designs  
 (B) cracks  
 (C) heads  
 (D) carvings
49. According to the passage, Borglum’s son  
 (A) is dead  
 (B) was a president  
 (C) did maintenance work  
 (D) spent a million dollars
50. The men who Borglum hired were  
 (A) trained sculptors  
 (B) laid-off stone masons  
 (C) Black Hills volunteers  
 (D) unemployed miners
51. It can be inferred from the passage that  
 (A) the heads are not as originally planned  
 (B) the workers made mistakes when blasting  
 (C) the cracks caused serious injuries  
 (D) the designs had large fissures in them
52. Borglum’s mixture for filling cracks was  
 (A) very expensive  
 (B) bought at the Black Hills mines  
 (C) invented by the sculptor himself  
 (D) uncovered during carving
53. Today Mount Rushmore needs  
 (A) to be protected from air pollution  
 (B) to be polished for tourists  
 (C) to be restored during the winter  
 (D) to be repaired periodically

## Questions 54–60

- 1 History books record that the first film with sound was *The Jazz Singer* in 1927.  
2 But sound films, or “talkies,” did not suddenly appear after years of silent  
3 screenings. From the earliest public performances in 1896, films were  
4 accompanied by music and sound effects. These were produced by a single  
5 pianist, a small band, or a full-scale orchestra; large movie theaters could buy  
6 sound-effects machines. Research into sound that was reproduced at exactly the  
7 same time as the pictures – called “synchronized sound” – began soon after the  
8 very first films were shown. With synchronized sound, characters on the movie  
9 screen could sing and speak. As early as 1896, the newly invented gramophone,  
10 which played a large disc carrying music and dialogue, was used as a sound  
11 system. The biggest disadvantage was that the sound and pictures could become  
12 unsynchronized if, for example, the gramophone needle jumped or if the speed of  
13 the projector changed. This system was only effective for a single song or dialogue  
14 sequence.
- 15 In the “sound-on-film” system, sounds were recorded as a series of marks on  
16 celluloid which could be read by an optical sensor. These signals would be placed  
17 on the film alongside the image, guaranteeing synchronization. Short feature films  
18 were produced in this way as early as 1922. This system eventually brought us  
19 “talking pictures.”
54. The passage is mainly about  
(A) the history of silent movies  
(B) the disadvantages of synchronized sound  
(C) the development of sound with movies  
(D) the research into sound reproduction
55. According to the passage, films using sound effects were screened  
(A) before 1896  
(B) as early as 1896  
(C) as early as 1922  
(D) in 1927
56. It can be inferred that  
(A) most movie theaters had a pianist  
(B) sound-effects machines were not common because they were expensive  
(C) orchestras couldn't synchronize sound with the pictures  
(D) gramophones were developed about the same time as moving pictures
57. Which of the following is NOT mentioned as a producer of sound to accompany movies?  
(A) a jazz singer  
(B) a single pianist  
(C) a small band  
(D) a gramophone
58. According to the passage, gramophones were ineffective because they  
(A) got out of synchronization with the picture  
(B) were too large for most movie theaters  
(C) were newly invented and still had imperfections  
(D) changed speeds when the needle jumped

3 • 3 • 3 • 3 • 3 • 3 • 3 • 3

59. According to the passage, sound-on-film guaranteed synchronization because the recording was
- (A) made during the filming of the picture
  - (B) read by an optical sensor
  - (C) placed beside the image on the film
  - (D) marked on the gramophone
60. Short feature films produced as early as 1922
- (A) were recorded by optical sensors
  - (B) put musicians out of work
  - (C) were only effective for dialogue sequences
  - (D) preceded talking pictures

**THIS IS THE END OF SECTION 3 OF PRACTICE TEST 2.**

IF YOU FINISH BEFORE TIME IS UP, CHECK YOUR WORK  
ON SECTION 3 ONLY.  
DO NOT READ OR WORK ON ANY OTHER SECTION OF THE TEST.



# ANSWER SHEET FOR PRACTICE TEST 2

## SECTION 1

1. A B C D
2. A B C D
3. A B C D
4. A B C D
5. A B C D
6. A B C D
7. A B C D
8. A B C D
9. A B C D
10. A B C D
11. A B C D
12. A B C D
13. A B C D
14. A B C D
15. A B C D
16. A B C D
17. A B C D
18. A B C D
19. A B C D
20. A B C D
21. A B C D
22. A B C D
23. A B C D
24. A B C D
25. A B C D
26. A B C D
27. A B C D
28. A B C D
29. A B C D
30. A B C D
31. A B C D
32. A B C D
33. A B C D
34. A B C D
35. A B C D
36. A B C D
37. A B C D
38. A B C D
39. A B C D
40. A B C D
41. A B C D
42. A B C D
43. A B C D
44. A B C D
45. A B C D
46. A B C D
47. A B C D
48. A B C D
49. A B C D
50. A B C D

## SECTION 2

1. A B C D
2. A B C D
3. A B C D
4. A B C D
5. A B C D
6. A B C D
7. A B C D
8. A B C D
9. A B C D
10. A B C D
11. A B C D
12. A B C D
13. A B C D
14. A B C D
15. A B C D
16. A B C D
17. A B C D
18. A B C D
19. A B C D
20. A B C D
21. A B C D
22. A B C D
23. A B C D
24. A B C D
25. A B C D
26. A B C D
27. A B C D
28. A B C D
29. A B C D
30. A B C D
31. A B C D
32. A B C D
33. A B C D
34. A B C D
35. A B C D
36. A B C D
37. A B C D
38. A B C D
39. A B C D
40. A B C D

## SECTION 3

- |             |             |
|-------------|-------------|
| 1. A B C D  | 31. A B C D |
| 2. A B C D  | 32. A B C D |
| 3. A B C D  | 33. A B C D |
| 4. A B C D  | 34. A B C D |
| 5. A B C D  | 35. A B C D |
| 6. A B C D  | 36. A B C D |
| 7. A B C D  | 37. A B C D |
| 8. A B C D  | 38. A B C D |
| 9. A B C D  | 39. A B C D |
| 10. A B C D | 40. A B C D |
| 11. A B C D | 41. A B C D |
| 12. A B C D | 42. A B C D |
| 13. A B C D | 43. A B C D |
| 14. A B C D | 44. A B C D |
| 15. A B C D | 45. A B C D |
| 16. A B C D | 46. A B C D |
| 17. A B C D | 47. A B C D |
| 18. A B C D | 48. A B C D |
| 19. A B C D | 49. A B C D |
| 20. A B C D | 50. A B C D |
| 21. A B C D | 51. A B C D |
| 22. A B C D | 52. A B C D |
| 23. A B C D | 53. A B C D |
| 24. A B C D | 54. A B C D |
| 25. A B C D | 55. A B C D |
| 26. A B C D | 56. A B C D |
| 27. A B C D | 57. A B C D |
| 28. A B C D | 58. A B C D |
| 29. A B C D | 59. A B C D |
| 30. A B C D | 60. A B C D |

DATOS DE LOS PARTICIPANTES EN LA COMPARACIÓN DE PRUEBAS TOEFL Y FCE

APÉNDICE 5

Participantes	Edad	Género	Años de Est.		Tipo de Inst.	Años Est.	Certific. Cambridge	TOEFLComp Auditiva	TOEFL Exp. Esc. Gram	TOEFLVoc. Comp. Lec.	TOTAL	FCE Comp.		FCE Gram.		FCE Comp.		FCE Exp.	TOTAL
			Inglés	Inglés								Lectura	Vocab.	Auditiva	Oral				
JR	18	0	14	0				28	15	28	71	31	35	21				87	
EF	17	0	14	0	0	9		34	27	38	99	27	41	20	24			112	
MT	17	0	13	0	2		1	46	31	47	124	35	57	24	25			141	
MP	19	1	13	0	0	10		39	19	33	91	31	41	20	19			111	
IN	17	0	12	0				33	21	35	89	30	42	22				94	
RS	18	0	15	0				32	20	44	96	29	38	19				86	
AD	18	0	12	0			1	36	24	40	100	32	53	24	23			132	
KL	18	0	14	0	0	14		23	17	29	69	28	46	13				87	
JM	17	0	14	0				29	16	27	72	26	36	21				83	
AV	17	1	12	0	0	12		38	20	42	100	31	47	21				99	
PY	17	0	12	0			1	36	20	35	91	31	46	23	23			123	
RA	17	1	13	0				30	18	30	78	25	39	17				81	
JG	17	1	12	0			1	32	21	29	82	27	37	17				81	
HF	18	1	12	0	1	1		29	26	33	88	21	43	22				86	
BZ	17	1	12	0	2			35	14	38	87	28	41	22	22			113	
MS	18	0	6	1	0	3	1	30	21	34	85	35	43	23				101	
LH	18	0	5	1	1	2	2	42	20	38	100	28	43	16	24			111	
TJ	19	1	13	0	0	9		30	18	42	90	22	32	19				73	
BD	17	1	11	1				33	18	27	78	29	30	17	19			95	
SU	17	0	11	0	3	11		47	32	50	129	35	57	24				116	
CO	18	1	14	0	2		1	38	12	27	77	26	44	27	19			116	
VA	19	1	15	0	3	4	1	29	17	34	80	24	28	16				68	
OJ	18	1	15	0	2		1	34	26	48	108	32	55	25	24			136	
A	B	C	D	E	F	G	H	I	J	K		L	M	N	.	O			

A. Participantes

B. Edad

C. Género: F= 0 M= 1

D. Años de estudio de inglés

E. Estudió inglés en:kinder, prim. sec. prep.0, secundaria y preparatoria 1

F. Tipo de Institución: escuelas bilingües 0, de idiomas 1,dos o más cursos de verano en el extranjero 2, prim. y sec en EUA, 3

- G. Años de estudio de inglés.
- H. Certificaciones Cambridge: KET 0, PET 1, FCE 2
- I. TOEFL Comprensión Auditiva
- J. TOEFL Expresión Escrita y Gramática
- K. TOEFL Vocabulario y Comprensión de Lectura
- L. FCE Comprensión de Lectura
- M. FCE Gramática y Vocabulario
- N. FCE Comprensión Auditiva
- O. FCE Expresión Oral